PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/86, 15/35, C07K 14/015, C12N 5/10, C07K 16/08, G01N 33/53, C12Q 1/68, A61K 48/00

A2

(11) International Publication Number: WO 98/11244

(43) International Publication Date: 19 March 1998 (19.03.98)

(21) International Application Number: PCT/US97/16266

(22) International Filing Date: 11 September 1997 (11.09.97)

(30) Priority Data:
60/025,934 11 September 1996 (11.09.96) US

(71) Applicant (*for all designated States except US*): THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; National Institutes of Health, Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).

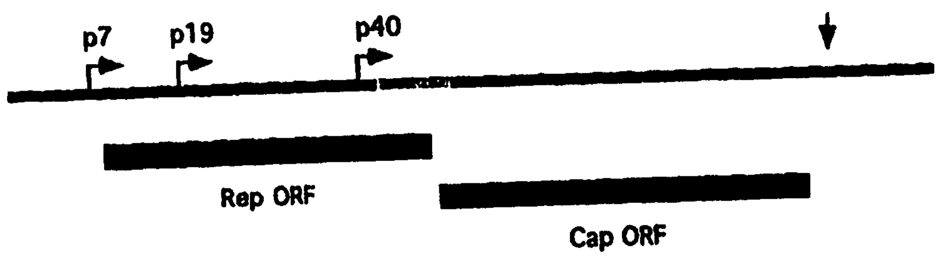
(72) Inventors; and
(75) Inventors/Applicants (*for US only*): CHIORINI, John, A. [US/US]; 2604 Loma Street, Silver Spring, MD 20902 (US). KOTIN, Robert, M. [US/US]; 707 Gormley, Rockville, MD 20850 (US). SAFER, Brian [US/US]; 1610 Tilton Drive, Silver Springs, MD 20902 (US).

(74) Agents: SELBY, Elizabeth et al.; Needle & Rosenberg, Suite 1200, 127 Peachtree Street, N.E., Atlanta, GA 30303 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: AAV4 VECTOR AND USES THEREOF



(57) Abstract

The present invention provides an adeno-associated virus 4 (AAV4) virus and vectors and particles derived therefrom. In addition, the present invention provides methods of delivering a nucleic acid to a cell using the AAV4 vectors and particles.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece		Republic of Macedonia	**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's	**NZ**	New Zealand		
CM	Cameroon		Republic of Korea	**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

AAV4 VECTOR AND USES THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention provides adeno-associated virus 4 (AAV4) and vectors derived therefrom. Thus, the present invention relates to AAV4 vectors for and methods of delivering nucleic acids to cells of subjects.

Background Art

Adeno associated virus (AAV) is a small nonpathogenic virus of the parvoviridae family (for review see 28). AAV is distinct from the other members of this family by its dependence upon a helper virus for replication. In the absence of a helper virus, AAV may integrate in a locus specific manner into the q arm of chromosome 19 (21). The approximately 5 kb genome of AAV consists of one segment of single stranded DNA of either plus or minus polarity. The ends of the genome are short inverted terminal repeats which can fold into hairpin structures and serve as the origin of viral DNA replication. Physically, the parvovirus virion is non-enveloped and its icosohedral capsid is approximately 20 nm in diameter.

To date 7 serologically distinct AAVs have been identified and 5 have been isolated from humans or primates and are referred to as AAV types 1-5 (1). The most extensively studied of these isolates is AAV type 2 (AAV2). The genome of AAV2 is 4680 nucleotides in length and contains two open reading frames (ORFs). The left ORF encodes the non-structural Rep proteins, Rep40, Rep 52, Rep68 and Rep 78, which are involved in regulation of replication and transcription in addition to the production of single-stranded progeny genomes (5-8, 11, 12, 15, 17, 19, 21-23, 25, 34, 37-40). Furthermore, two of the Rep proteins have been associated with the preferential

integration of AAV genomes into a region of the q arm of human chromosome 19. Rep68/78 have also been shown to possess NTP binding activity as well as DNA and RNA helicase activities. The Rep proteins possess a nuclear localization signal as well as several potential phosphorylation sites. Mutation of one of these kinase sites resulted in a loss of replication activity.

The ends of the genome are short inverted terminal repeats which have the potential to fold into T-shaped hairpin structures that serve as the origin of viral DNA replication. Within the ITR region two elements have been described which are central to the function of the ITR, a GAGC repeat motif and the terminal resolution site (trs). The repeat motif has been shown to bind Rep when the ITR is in either a linear or hairpin conformation (7, 8, 26). This binding serves to position Rep68/78 for cleavage at the trs which occurs in a site- and strand-specific manner. In addition to their role in replication, these two elements appear to be central to viral integration. Contained within the chromosome 19 integration locus is a Rep binding site with an adjacent trs. These elements have been shown to be functional and necessary for locus specific integration.

The AAV2 virion is a non-enveloped, icosohedral particle approximately 25 nm in diameter, consisting of three related proteins referred to as VPI,2 and 3. The right ORF encodes the capsid proteins, VP1, VP2, and VP3. These proteins are found in a ratio of 1:1:10 respectively and are all derived from the right-hand ORF. The capsid proteins differ from each other by the use of alternative splicing and an unusual start codon. Deletion analysis has shown that removal or alteration of VPl which is translated from an alternatively spliced message results in a reduced yield of infections particles (15, 16, 38). Mutations within the VP3 coding region result in the failure to produce any single-stranded progeny DNA or infectious particles (15, 16, 38).

The following features of AAV have made it an attractive vector for gene transfer (16). AAV vectors have been shown *in vitro* to stably integrate into the cellular genome; possess a broad host range; transduce both dividing and non dividing cells *in vitro* and *in vivo* (13, 20, 30, 32) and maintain high levels of expression of the transduced genes (41). Viral particles are heat stable, resistant to solvents, detergents, changes in pH, temperature, and can be concentrated on CsCl gradients (1,2).

Integration of AAV provirus is not associated with any long term negative effects on cell growth or differentiation (3,42). The ITRs have been shown to be the only *cis* elements required for replication, packaging and integration (35) and may contain some promoter activities (14).

Initial data indicate that AAV4 is a unique member of this family. DNA hybridization data indicated a similar level of homology for AAV1-4 (31). However, in contrast to the other AAVs only one ORF corresponding to the capsid proteins was identified in AAV4 and no ORF was detected for the Rep proteins (27).

AAV2 was originally thought to infect a wide variety of cell types provided the appropriate helper virus was present. Recent work has shown that some cell lines are transduced very poorly by AAV2 (30). While the receptor has not been completely characterized, binding studies have indicated that it is poorly expressed on erythroid cells (26). Recombinant AAV2 transduction of $CD34^+$, bone marrow pluripotent cells, requires a multiplicity of infection (MOI) of 10^4 particles per cell (A. W. Nienhuis unpublished results). This suggests that transduction is occurring by a non-specific mechanism or that the density of receptors displayed on the cell surface is low compared to other cell types.

The present invention provides a vector comprising the AAV4 virus as well as AAV4 viral particles. While AAV4 is similar to AAV2, the two viruses are found herein to be physically and genetically distinct. These differences endow AAV4 with some unique advantages which better suit it as a vector for gene therapy. For example, the wt AAV4 genome is larger than AAV2, allowing for efficient encapsidation of a larger recombinant genome. Furthermore, wt AAV4 particles have a greater buoyant density than AAV2 particles and therefore are more easily separated from contaminating helper virus and empty AAV particles than AAV2-based particles. Additionally, in contrast to AAV1, 2, and 3, AAV4, is able to hemagglutinate human, guinea pig, and sheep erythrocytes (18).

Furthermore, as shown herein, AAV4 capsid protein, again surprisingly, is distinct from AAV2 capsid protein and exhibits different tissue tropism. AAV2 and AAV4 have been shown to be serologically distinct and thus, in a gene therapy application, AAV4 would allow for transduction of a patient who already possess neutralizing antibodies to AAV2 either as a result of natural immunological defense or from prior exposure to AAV2 vectors. Thus, the present invention, by providing these new recombinant vectors and particles based on AAV4 provides a new and highly useful series of vectors.

SUMMARY OF THE INVENTION

The present invention provides a nucleic acid vector comprising a pair of adeno-associated virus 4 (AAV4) inverted terminal repeats and a promoter between the inverted terminal repeats.

The present invention further provides an AAV4 particle containing a vector comprising a pair of AAV2 inverted terminal repeats.

Additionally, the instant invention provides an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 [AAV4 genome]. Furthermore, the present invention provides an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:1 [AAV4 genome].

The present invention provides an isolated nucleic acid encoding an adeno-associated virus 4 Rep protein. Additionally provided is an isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:2, or a unique fragment thereof. Additionally provided is an isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:8, or a unique fragment thereof. Additionally provided is an isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:9, or a unique fragment thereof. Additionally provided is an isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:10, or a unique fragment thereof. Additionally provided is an isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:11, or a unique fragment thereof.

The present invention further provides an isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:4. Additionally provided is an isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:16. Also provided is an isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:18.

The present invention additionally provides an isolated nucleic acid encoding adeno-associated virus 4 capsid protein.

The present invention further provides an AAV4 particle comprising a capsid protein consisting essentially of the amino acid sequence set forth in SEQ ID NO:4.

Additionally provided by the present invention is an isolated nucleic acid comprising an AAV4 p5 promoter.

The instant invention provides a method of screening a cell for infectivity by AAV4 comprising contacting the cell with AAV4 and detecting the presence of AAV4 in the cells.

The present invention further provides a method of delivering a nucleic acid to a cell comprising administering to the cell an AAV4 particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell.

The present invention also provides a method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

The present invention further provides a method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

The present invention also provides a method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV4 particle comprising

the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject.

The instant invention further provides a method of delivering a nucleic acid to a cell in a subject having antibodies to AAV2 comprising administering to the subject an AAV4 particle comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows a schematic outline of AAV 4. Promoters are indicated by horizontal arrows with their corresponding map positions indicated above. The polyadenylation site is indicated by a vertical arrow and the two open reading frames are indicated by black boxes. The splice region is indicated by a shaded box.

Fig. 2 shows AAV4 ITR.The sequence of the ITR (SEQ ID NO: 20) is shown in the hairpin conformation. The putative Rep binding site is boxed. The cleavage site in the trs is indicated by an arrow. Bases which differ from the ITR of AAV2 are outlined.

Fig. 3 shows cotransduction of rAAV2 and rAAV4. Cos cells were transduced with a constant amount of rAAV2 or rAAV4 expressing beta galactosidase and increasing amounts of rAAV2 expressing human factor IX (rAAV2FIX) . For the competition the number of positive cells detected in the cotransduced wells was divided by the number of positive cells in the control wells (cells transduced with only rAAV2LacZ or rAAV4LacZ) and expressed as a percent of the control. This value was plotted against the number of particles of rAAV2FIX.

Fig. 4 shows effect of trypsin treatment on cos cell transduction. Cos cell monolayers were trypsinized and diluted in complete media. Cells were incubated with virus at an MOI of 260 and following cell attachment the virus was removed. As a control an equal number of cos cells were plated and allowed to attach overnight before transduction with virus for the same amount of time. The number of positive cells was determined by staining 50 hrs post transduction. The data is presented as a ratio of the number of positive cells seen with the trypsinized group and the control group.

DETAILED DESCRIPTION OF THE INVENTION

As used in the specification and in the claims, "a" can mean one or more, depending upon the context in which it is used.

The present invention provides the nucleotide sequence of the adeno-associated virus 4 (AAV4) genome and vectors and particles derived therefrom. Specifically, the present invention provides a nucleic acid vector comprising a pair of AAV4 inverted terminal repeats (ITRs) and a promoter between the inverted terminal repeats. The AAV4 ITRs are exemplified by the nucleotide sequence set forth in SEQ ID NO:6 and SEQ ID NO:20; however, these sequences can have minor modifications and still be contemplated to constitute AAV4 ITRs. The nucleic acid listed in SEQ ID NO:6 depicts the ITR in the "flip" orientation of the ITR. The nucleic acid listed in SEQ ID NO:20 depicts the ITR in the "flop" orientation of the ITR. Minor modifications in an ITR of either orientation are those that will not interfere with the hairpin structure formed by the AAV4 ITR as described herein and known in the art. Furthermore, to be considered within the term "AAV4 ITRs" the nucleotide sequence must retain the Rep binding site described herein and exemplified in SEQ ID NO:6 and SEQ ID NO:20, *i.e.*, it must retain one or both features described herein that distinguish the AAV4 ITR from the AAV2 ITR: (1) four (rather than three as in AAV2) "GAGC" repeats and (2) in the AAV4 ITR Rep binding site the fourth nucleotide in the first two "GAGC" repeats is a T rather than a C.

The promoter can be any desired promoter, selected by known considerations, such as the level of expression of a nucleic acid functionally linked to the promoter and the cell type in which the vector is to be used. Promoters can be an exogenous or an endogenous promoter. Promoters can include, for example, known strong promoters such as SV40 or the inducible metallothionein promoter, or an AAV promoter, such as an AAV p5 promoter. Additional examples of promoters include promoters derived from actin genes, immunoglobulin genes, cytomegalovirus (CMV), adenovirus, bovine

papilloma virus, adenoviral promoters, such as the adenoviral major late promoter, an inducible heat shock promoter, respiratory syncytial virus, Rous sarcomas virus (RSV), etc. Specifically, the promoter can be AAV2 p5 promoter or AAV4 p5 promoter. More specifically, the AAV4 p5 promoter can be about nucleotides 130 to 291 of SEQ ID NO: 1. Additionally, the p5 promoter may be enhanced by nucleotides 1-130. Furthermore, smaller fragments of p5 promoter that retain promoter activity can readily be determined by standard procedures including, for example, constructing a series of deletions in the p5 promoter, linking the deletion to a reporter gene, and determining whether the reporter gene is expressed, *i.e.*, transcribed and/or translated.

It should be recognized that the nucleotide and amino acid sequences set forth herein may contain minor sequencing errors. Such errors in the nucleotide sequences can be corrected, for example, by using the hybridization procedure described above with various probes derived from the described sequences such that the coding sequence can be reisolated and resequenced. The corresponding amino acid sequence can then be corrected accordingly.

The AAV4 vector can further comprise an exogenous nucleic acid functionally linked to the promoter. By "heterologous nucleic acid" is meant that any heterologous or exogenous nucleic acid can be inserted into the vector for transfer into a cell, tissue or organism. The nucleic acid can encode a polypeptide or protein or an antisense RNA, for example. By "functionally linked" is meant such that the promoter can promote expression of the heterologous nucleic acid, as is known in the art, such as appropriate orientation of the promoter relative to the heterologous nucleic acid. Furthermore, the heterologous nucleic acid preferably has all appropriate sequences for expression of the nucleic acid, as known in the art, to functionally encode, *i.e.*, allow the nucleic acid to be expressed. The nucleic acid can include, for example, expression control sequences, such as an enhancer, and necessary information processing sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional terminator sequences.

The heterologous nucleic acid can encode beneficial proteins that replace missing or defective proteins required by the subject into which the vector in transferred or can encode a cytotoxic polypeptide that can be directed, *e.g.*, to cancer cells or other cells whose death would be beneficial to the subject. The heterologous nucleic acid can also encode antisense RNAs that can bind to, and thereby inactivate, mRNAs made by the subject that encode harmful proteins. In one embodiment, antisense polynucleotides can be produced from a heterologous expression cassette in an AAV4 viral construct where the expression cassette contains a sequence that promotes cell-type specific expression (Wirak *et al.*, *EMBO* 10:289 (1991)). For general methods relating to antisense polynucleotides, see *Antisense RNA and DNA*, D. A. Melton, Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988).

Examples of heterologous nucleic acids which can be administered to a cell or subject as part of the present AAV4 vector can include, but are not limited to the following: nucleic acids encoding therapeutic agents, such as tumor necrosis factors (TNF), such as TNF-α; interferons, such as interferon-α, interferon-β, and interferon-γ; interleukins, such as IL-1, IL-1β, and ILs -2 through -14; GM-CSF; adenosine deaminase; cellular growth factors, such as lymphokines; soluble CD4; Factor VIII; Factor IX; T-cell receptors; LDL receptor; ApoE; ApoC; alpha-1 antitrypsin; ornithine transcarbamylase (OTC); cystic fibrosis transmembrane receptor (CFTR); insulin; Fc receptors for antigen binding domains of antibodies, such as immunoglobulins; and antisense sequences which inhibit viral replication, such as antisense sequences which inhibit replication of hepatitis B or hepatitis non-A, non-B virus. The nucleic acid is chosen considering several factors, including the cell to be transfected. Where the target cell is a blood cell, for example, particularly useful nucleic acids to use are those which allow the blood cells to exert a therapeutic effect, such as a gene encoding a clotting factor for use in treatment of hemophilia. Furthermore, the nucleic acid can encode more than one gene product, limited only, if the nucleic acid is to be packaged in a capsid, by the size of nucleic acid that can be packaged.

Furthermore, suitable nucleic acids can include those that, when transferred into a primary cell, such as a blood cell, cause the transferred cell to target a site in the body where that cell's presence would be beneficial. For example, blood cells such as TIL cells can be modified, such as by transfer into the cell of a Fab portion of a monoclonal antibody, to recognize a selected antigen. Another example would be to introduce a nucleic acid that would target a therapeutic blood cell to tumor cells. Nucleic acids useful in treating cancer cells include those encoding chemotactic factors which cause an inflammatory response at a specific site, thereby having a therapeutic effect.

Cells, particularly blood cells, having such nucleic acids transferred into them can be useful in a variety of diseases, syndromes and conditions. For example, suitable nucleic acids include nucleic acids encoding soluble CD4, used in the treatment of AIDS and α-antitrypsin, used in the treatment of emphysema caused by α-antitrypsin deficiency. Other diseases, syndromes and conditions in which such cells can be useful include, for example, adenosine deaminase deficiency, sickle cell deficiency, brain disorders such as Alzheimer's disease, thalassemia, hemophilia, diabetes, phenylketonuria, growth disorders and heart diseases, such as those caused by alterations in cholesterol metabolism, and defects of the immune system.

As another example, hepatocytes can be transfected with the present vectors having useful nucleic acids to treat liver disease. For example, a nucleic acid encoding OTC can be used to transfect hepatocytes (*ex vivo* and returned to the liver or *in vivo*) to treat congenital hyperammonemia, caused by an inherited deficiency in OTC. Another example is to use a nucleic acid encoding LDL to target hepatocytes *ex vivo* or *in vivo* to treat inherited LDL receptor deficiency. Such transfected hepatocytes can also be used to treat acquired infectious diseases, such as diseases resulting from a viral infection. For example, transduced hepatocyte precursors can be used to treat viral hepatitis, such as hepatitis B and non-A, non-B hepatitis, for example by transducing the hepatocyte precursor with a nucleic acid encoding an antisense RNA that inhibits viral replication. Another example includes transferring a vector of the present invention

having a nucleic acid encoding a protein, such as α-interferon, which can confer resistance to the hepatitis virus.

For a procedure using transfected hepatocytes or hepatocyte precursors, hepatocyte precursors having a vector of the present invention transferred in can be grown in tissue culture, removed form the tissue culture vessel, and introduced to the body, such as by a surgical method. In this example, the tissue would be placed directly into the liver, or into the body cavity in proximity to the liver, as in a transplant or graft. Alternatively, the cells can simply be directly injected into the liver, into the portal circulatory system, or into the spleen, from which the cells can be transported to the liver via the circulatory system. Furthermore, the cells can be attached to a support, such as microcarrier beads, which can then be introduced, such as by injection, into the peritoneal cavity. Once the cells are in the liver, by whatever means, the cells can then express the nucleic acid and/or differentiate into mature hepatocytes which can express the nucleic acid.

The present invention also contemplates any unique fragment of these AAV4 nucleic acids, including the AAV4 nucleic acids set forth in SEQ ID NOs: 1, 3, 5, 6, 7, 12-15, 17 and 19. To be unique, the fragment must be of sufficient size to distinguish it from other known sequences, most readily determined by comparing any nucleic acid fragment to the nucleotide sequences of nucleic acids in computer databases, such as GenBank. Such comparative searches are standard in the art. Typically, a unique fragment useful as a primer or probe will be at least about 8 or 10 to about 20 or 25 nucleotides in length, depending upon the specific nucleotide content of the sequence. Additionally, fragments can be, for example, at least about 30, 40, 50, 75, 100, 200 or 500 nucleotides in length. The nucleic acid can be single or double stranded, depending upon the purpose for which it is intended.

The present invention further provides an AAV4 capsid protein. In particular, the present invention provides not only a polypeptide comprising all three AAV4 coat proteins, *i.e.*, VP1, VP2 and VP3, but also a polypeptide comprising each AAV4 coat

protein individually. Thus an AAV4 particle comprising an AAV4 capsid protein comprises at least one AAV4 coat protein VP1, VP2 or VP3. An AAV4 particle comprising an AAV4 capsid protein can be utilized to deliver a nucleic acid vector to a cell, tissue or subject. For example, the herein described AAV4 vectors can be encapsulated in an AAV4 particle and utilized in a gene delivery method. Furthermore, other viral nucleic acids can be encapsidated in the AAV4 particle and utilized in such delivery methods. For example, an AAV2 vector can be encapsidated in an AAV4 particle and administered. Furthermore, a chimeric capsid protein incorporating both AAV2 and AAV4 sequences can be generated, by standard cloning methods, selecting regions from each protein as desired. For example, particularly antigenic regions of the AAV2 capsid protein can be replaced with the corresponding region of the AAV4 capsid protein.

The herein described AAV4 nucleic acid vector can be encapsidated in an AAV particle. In particular, it can be encapsidated in an AAV1 particle, an AAV2 particle, an AAV3 particle, an AAV4 particle, or an AAV5 particle by standard methods using the appropriate capsid proteins in the encapsidation process, as long as the nucleic acid vector fits within the size limitation of the particle utilized. The encapsidation process itself is standard in the art.

An AAV4 particle is a viral particle comprising an AAV4 capsid protein. An AAV4 capsid polypeptide encoding the entire VP1, VP2, and VP3 polypeptide can overall have at least about 63% homology to the polypeptide having the amino acid sequence encoded by nucleotides 2260-4464 set forth in SEQ ID NO:1 (AAV4 capsid protein). The capsid protein can have about 70% homology, about 75% homology, 80% homology, 85% homology, 90% homology, 95% homology, 98% homology, 99% homology, or even 100% homology to the protein having the amino acid sequence encoded by nucleotides 2260-4464 set forth in SEQ ID NO:1. The particle can be a particle comprising both AAV4 and AAV2 capsid protein, *i.e.*, a chimeric protein. Variations in the amino acid sequence of the AAV4 capsid protein are contemplated herein, as long as the resulting viral particle comprising the AAV4 capsid remains

antigenically or immunologically distinct from AAV2, as can be routinely determined by standard methods. Specifically, for example, ELISA and Western blots can be used to determine whether a viral particle is antigenically or immunologically distinct from AAV2. Furthermore, the AAV4 viral particle preferably retains tissue tropism distinction from AAV2, such as that exemplified in the examples herein, though an AAV4 chimeric particle comprising at least one AAV4 coat protein may have a different tissue tropism from that of an AAV4 particle consisting only of AAV4 coat proteins.

The invention further provides an AAV4 particle containing, *i.e.*, encapsidating, a vector comprising a pair of AAV2 inverted terminal repeats. The nucleotide sequence of AAV2 ITRs is known in the art. Furthermore, the particle can be a particle comprising both AAV4 and AAV2 capsid protein, *i.e.*, a chimeric protein. The vector encapsidated in the particle can further comprise an exogenous nucleic acid inserted between the inverted terminal repeats.

The present invention further provides an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 (AAV4 genome). This nucleic acid, or portions thereof, can be inserted into other vectors, such as plasmids, yeast artificial chromosomes, or other viral vectors, if desired, by standard cloning methods. The present invention also provides an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:1. The nucleotides of SEQ ID NO:1 can have minor modifications and still be contemplated by the present invention. For example, modifications that do not alter the amino acid encoded by any given codon (such as by modification of the third, "wobble," position in a codon) can readily be made, and such alterations are known in the art. Furthermore, modifications that cause a resulting neutral amino acid substitution of a similar amino acid can be made in a coding region of the genome. Additionally, modifications as described herein for the AAV4 components, such as the ITRs, the p5 promoter, etc. are contemplated in this invention.

The present invention additionally provides an isolated nucleic acid that selectively hybridizes with an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:1 (AAV4 genome). The present invention further provides an isolated nucleic acid that selectively hybridizes with an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 (AAV4 genome). By "selectively hybridizes" as used in the claims is meant a nucleic acid that specifically hybridizes to the particular target nucleic acid under sufficient stringency conditions to selectively hybridize to the target nucleic acid without significant background hybridization to a nucleic acid encoding an unrelated protein, and particularly, without detectably hybridizing to AAV2. Thus, a nucleic acid that selectively hybridizes with a nucleic acid of the present invention will not selectively hybridize under stringent conditions with a nucleic acid encoding a different protein, and vice versa. Therefore, nucleic acids for use, for example, as primers and probes to detect or amplify the target nucleic acids are contemplated herein. Nucleic acid fragments that selectively hybridize to any given nucleic acid can be used, *e.g.*, as primers and or probes for further hybridization or for amplification methods (*e.g.*, polymerase chain reaction (PCR), ligase chain reaction (LCR)). Additionally, for example, a primer or probe can be designed that selectively hybridizes with both AAV4 and a gene of interest carried within the AAV4 vector (*i.e.*, a chimeric nucleic acid).

Stringency of hybridization is controlled by both temperature and salt concentration of either or both of the hybridization and washing steps. Typically, the stringency of hybridization to achieve selective hybridization involves hybridization in high ionic strength solution (6X SSC or 6X SSPE) at a temperature that is about 12-25°C below the T_m (the melting temperature at which half of the molecules dissociate from its partner) followed by washing at a combination of temperature and salt concentration chosen so that the washing temperature is about 5°C to 20°C below the T_m The temperature and salt conditions are readily determined empirically in preliminary experiments in which samples of reference DNA immobilized on filters are hybridized to a labeled nucleic acid of interest and then washed under conditions of different stringencies. Hybridization temperatures are typically higher for DNA-RNA and RNA-

RNA hybridizations. The washing temperatures can be used as described above to achieve selective stringency, as is known in the art. (Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989; Kunkel et al. *Methods Enzymol.* 1987:154:367, 1987). A preferable stringent hybridization condition for a DNA:DNA hybridization can be at about 68°C (in aqueous solution) in 6X SSC or 6X SSPE followed by washing at 68°C. Stringency of hybridization and washing, if desired, can be reduced accordingly as homology desired is decreased, and further, depending upon the G-C or A-T richness of any area wherein variability is searched for. Likewise, stringency of hybridization and washing, if desired, can be increased accordingly as homology desired is increased, and further, depending upon the G-C or A-T richness of any area wherein high homology is desired, all as known in the art.

A nucleic acid that selectively hybridizes to any portion of the AAV4 genome is contemplated herein. Therefore, a nucleic acid that selectively hybridizes to AAV4 can be of longer length than the AAV4 genome, it can be about the same length as the AAV4 genome or it can be shorter than the AAV4 genome. The length of the nucleic acid is limited on the shorter end of the size range only by its specificity for hybridization to AAV4, *i.e.*, once it is too short, typically less than about 5 to 7 nucleotides in length, it will no longer bind specifically to AAV4, but rather will hybridize to numerous background nucleic acids. Additionally contemplated by this invention is a nucleic acid that has a portion that specifically hybridizes to AAV4 and a portion that specifically hybridizes to a gene of interest inserted within AAV4.

The present invention further provides an isolated nucleic acid encoding an adeno-associated virus 4 Rep protein. The AAV4 Rep proteins are encoded by open reading frame (ORF) 1 of the AAV4 genome. The AAV4 Rep genes are exemplified by the nucleic acid set forth in SEQ ID NO:3 (AAV4 ORF1), and include a nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:3 and a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:3. The present invention also includes a nucleic acid encoding the amino acid sequence set forth in SEQ

ID NO: 2 (polypeptide encoded by AAV4 ORF1). However, the present invention includes that the Rep genes nucleic acid can include any one, two, three, or four of the four Rep proteins, in any order, in such a nucleic acid. Furthermore, minor modifications are contemplated in the nucleic acid, such as silent mutations in the coding sequences, mutations that make neutral or conservative changes in the encoded amino acid sequence, and mutations in regulatory regions that do not disrupt the expression of the gene. Examples of other minor modifications are known in the art. Further modifications can be made in the nucleic acid, such as to disrupt or alter expression of one or more of the Rep proteins in order to, for example, determine the effect of such a disruption; such as to mutate one or more of the Rep proteins to determine the resulting effect, etc. However, in general, a modified nucleic acid encoding all four Rep proteins will have at least about 90%, about 93%, about 95%, about 98% or 100% homology to the sequence set forth in SEQ ID NO:3, and the Rep polypeptide encoded therein will have overall about 93%, about 95%, about 98%, about 99% or 100% homology with the amino acid sequence set forth in SEQ ID NO:2.

The present invention also provides an isolated nucleic acid that selectively hybridizes with a nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:3 and an isolated nucleic acid that selectively hybridizes with a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:3. "Selectively hybridizing" is defined elsewhere herein.

The present invention also provides each individual AAV4 Rep protein and the nucleic acid encoding each. Thus the present invention provides the nucleic acid encoding a Rep 40 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:12, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:12, and a nucleic acid encoding the adeno-associated virus 4 Rep protein having the amino acid sequence set forth in SEQ ID NO:8. The present invention also provides the nucleic acid encoding a Rep 52 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:13, an isolated nucleic acid consisting essentially of

the nucleotide sequence set forth in SEQ ID NO:13, and a nucleic acid encoding the adeno-associated virus 4 Rep protein having the amino acid sequence set forth in SEQ ID NO:9. The present invention further provides the nucleic acid encoding a Rep 68 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:14, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:14, and a nucleic acid encoding the adeno-associated virus 4 Rep protein having the amino acid sequence set forth in SEQ ID NO:10. And, further, the present invention provides the nucleic acid encoding a Rep 78 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:15, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:15, and a nucleic acid encoding the adeno-associated virus 4 Rep protein having the amino acid sequence set forth in SEQ ID NO:11. As described elsewhere herein, these nucleic acids can have minor modifications, including silent nucleotide substitutions, mutations causing neutral amino acid substitutions in the encoded proteins, and mutations in control regions that do not or minimally affect the encoded amino acid sequence.

The present invention further provides a nucleic acid encoding the entire AAV4 Capsid polypeptide. Specifically, the present invention provides a nucleic acid having the nucleotide sequence set for the nucleotides 2260-4464 of SEQ ID NO:1. Furthermore, the present invention provides a nucleic acid encoding each of the three AAV4 coat proteins, VP1, VP2, and VP3. Thus, the present invention provides a nucleic acid encoding AAV4 VP1, a nucleic acid encoding AAV4 VP2, and a nucleic acid encoding AAV4 VP3. Thus, the present invention provides a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:4 (VP1); a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:16 (VP2), and a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:18 (VP3). The present invention also specifically provides a nucleic acid comprising SEQ ID NO:5 (VP1 gene); a nucleic acid comprising SEQ ID NO:17 (VP2 gene); and a nucleic acid comprising SEQ ID NO:19 (VP3 gene). The present invention also specifically provides a nucleic acid consisting essentially of SEQ ID NO:5 (VP1 gene), a nucleic acid consisting essentially of SEQ ID

NO:17 (VP2 gene), and a nucleic acid consisting essentially of SEQ ID NO:19 (VP3 gene). Furthermore, a nucleic acid encoding an AAV4 capsid protein VP1 is set forth as nucleotides 2157-4361 of SEQ ID NO:1; a nucleic acid encoding an AAV4 capsid protein VP2 is set forth as nucleotides 2565-4361 of SEQ ID NO:1; and a nucleic acid encoding an AAV4 capsid protein VP3 is set forth as nucleotides 2745-4361 of SEQ ID NO:1. Minor modifications in the nucleotide sequences encoding the capsid, or coat, proteins are contemplated, as described above for other AAV4 nucleic acids.

The present invention also provides a cell containing one or more of the herein described nucleic acids, such as the AAV4 genome, AAV4 ORF1 and ORF2, each AAV4 Rep protein gene, and each AAV4 capsid protein gene. Such a cell can be any desired cell and can be selected based upon the use intended. For example, cells can include human HeLa cells, cos cells, other human and mammalian cells and cell lines. Primary cultures as well as established cultures and cell lines can be used. Nucleic acids of the present invention can be delivered into cells by any selected means, in particular depending upon the target cells. Many delivery means are well-known in the art. For example, electroporation, calcium phosphate precipitation, microinjection, cationic or anionic liposomes, and liposomes in combination with a nuclear localization signal peptide for delivery to the nucleus can be utilized, as is known in the art. Additionally, if in a viral particle, the cells can simply be transfected with the particle by standard means known in the art for AAV transfection.

The term "polypeptide" as used herein refers to a polymer of amino acids and includes full-length proteins and fragments thereof. Thus, "protein," polypeptide," and "peptide" are often used interchangeably herein. Substitutions can be selected by known parameters to be neutral (*see, e.g.*, Robinson WE Jr, and Mitchell WM., AIDS 4:S151-S162 (1990)). As will be appreciated by those skilled in the art, the invention also includes those polypeptides having slight variations in amino acid sequences or other properties. Such variations may arise naturally as allelic variations (*e.g.*, due to genetic polymorphism) or may be produced by human intervention (*e.g.*, by mutagenesis of cloned DNA sequences), such as induced point, deletion, insertion and substitution

mutants. Minor changes in amino acid sequence are generally preferred, such as conservative amino acid replacements, small internal deletions or insertions, and additions or deletions at the ends of the molecules. Substitutions may be designed based on, for example, the model of Dayhoff, *et al.* (in *Atlas of Protein Sequence and Structure 1978*, Nat'l Biomed. Res. Found., Washington, D.C.). These modifications can result in changes in the amino acid sequence, provide silent mutations, modify a restriction site, or provide other specific mutations.

A polypeptide of the present invention can be readily obtained by any of several means. For example, polypeptide of interest can be synthesized mechanically by standard methods. Additionally, the coding regions of the genes can be expressed and the resulting polypeptide isolated by standard methods. Furthermore, an antibody specific for the resulting polypeptide can be raised by standard methods (see, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988), and the protein can be isolated from a cell expressing the nucleic acid encoding the polypeptide by selective hybridization with the antibody. This protein can be purified to the extent desired by standard methods of protein purification (see, *e.g.*, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).

Typically, to be unique, a polypeptide fragment of the present invention will be at least about 5 amino acids in length; however, unique fragments can be 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more amino acids in length. A unique polypeptide will typically comprise such a unique fragment; however, a unique polypeptide can also be determined by its overall homology. A unique polypeptide can be 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more amino acids in length. Uniqueness of a polypeptide fragment can readily be determined by standard methods such as searches of computer databases of known peptide or nucleic acid sequences or by hybridization studies to the nucleic acid encoding the protein or to the protein itself, as known in the art.

The present invention provides an isolated AAV4 Rep protein. AAV4 Rep polypeptide is encoded by ORF1 of AAV4. Specifically, the present invention provides an AAV4 Rep polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, or a unique fragment thereof. The present invention also provides an AAV4 Rep polypeptide consisting essentially of the amino acid sequence set forth in SEQ ID NO:2, or a unique fragment thereof. Additionally, nucleotides 291-2306 of the AAV4 genome, which genome is set forth in SEQ ID NO:1, encode the AAV4 Rep polypeptide. The present invention also provides each AAV4 Rep protein. Thus the present invention provides AAV4 Rep 40, or a unique fragment thereof. The present invention particularly provides Rep 40 having the amino acid sequence set forth in SEQ ID NO:8. The present invention provides AAV4 Rep 52, or a unique fragment thereof. The present invention particularly provides Rep 52 having the amino acid sequence set forth in SEQ ID NO:9. The present invention provides AAV4 Rep 68, or a unique fragment thereof. The present invention particularly provides Rep 68 having the amino acid sequence set forth in SEQ ID NO:10. The present invention provides AAV4 Rep 78, or a unique fragment thereof. The present invention particularly provides Rep 78 having the amino acid sequence set forth in SEQ ID NO:11. By "unique fragment thereof" is meant any smaller polypeptide fragment encoded by AAV rep gene that is of sufficient length to be unique to the Rep polypeptide. Substitutions and modifications of the amino acid sequence can be made as described above and, further, can include protein processing modifications, such as glycosylation, to the polypeptide. However, a polypeptide including all four Rep proteins will encode a polypeptide having at least about 91% overall homology to the sequence set forth in SEQ ID NO:2, and it can have about 93%, about 95%, about 98%, about 99% or 100% homology with the amino acid sequence set forth in SEQ ID NO:2.

The present invention further provides an AAV4 Capsid polypeptide or a unique fragment thereof. AAV4 capsid polypeptide is encoded by ORF 2 of AAV4. Specifically, the present invention provides an AAV4 Capsid protein comprising the

amino acid sequence encoded by nucleotides 2260-4464 of the nucleotide sequence set forth in SEQ ID NO:1, or a unique fragment of such protein. The present invention also provides an AAV4 Capsid protein consisting essentially of the amino acid sequence encoded by nucleotides 2260-4464 of the nucleotide sequence set forth in SEQ ID NO:1, or a unique fragment of such protein. The present invention further provides the individual AAV4 coat proteins, VP1, VP2 and VP3. Thus, the present invention provides an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:4 (VP1). The present invention additionally provides an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:16 (VP2). The present invention also provides an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:18 (VP3). By "unique fragment thereof" is meant any smaller polypeptide fragment encoded by any AAV4 capsid gene that is of sufficient length to be unique to the AAV4 Capsid protein. Substitutions and modifications of the amino acid sequence can be made as described above and, further, can include protein processing modifications, such as glycosylation, to the polypeptide. However, an AAV4 Capsid polypeptide including all three coat proteins will have at least about 63% overall homology to the polypeptide encoded by nucleotides 2260-4464 of the sequence set forth in SEQ ID NO: 1. The protein can have about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95% or even 100% homology to the amino acid sequence encoded by the nucleotides 2260-4464 of the sequence set forth in SEQ ID NO:4. An AAV4 VP2 polypeptide can have at least about 58%, about 60%, about 70%, about 80%, about 90% about 95% or about 100% homology to the amino acid sequence set forth in SEQ ID NO:16. An AAV4 VP3 polypeptide can have at least about 60%, about 70%, about 80%, about 90% about 95% or about 100% homology to the amino acid sequence set forth in SEQ ID NO:18.

The present invention further provides an isolated antibody that specifically binds AAV4 Rep protein. Also provided is an isolated antibody that specifically binds the AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:2, or that specifically binds a unique fragment thereof. Clearly, any given antibody can recognize and bind one of a number of possible epitopes present in the polypeptide; thus only a

unique portion of a polypeptide (having the epitope) may need to be present in an assay to determine if the antibody specifically binds the polypeptide.

The present invention additionally provides an isolated antibody that specifically binds any adeno-associated virus 4 Capsid protein or the polypeptide comprising all three AAV4 coat proteins. Also provided is an isolated antibody that specifically binds the AAV4 Capsid protein having the amino acid sequence set forth in SEQ ID NO:4, or that specifically binds a unique fragment thereof. The present invention further provides an isolated antibody that specifically binds the AAV4 Capsid protein having the amino acid sequence set forth in SEQ ID NO:16, or that specifically binds a unique fragment thereof. The invention additionally provides an isolated antibody that specifically binds the AAV4 Capsid protein having the amino acid sequence set forth in SEQ ID NO:18, or that specifically binds a unique fragment thereof. Again, any given antibody can recognize and bind one of a number of possible epitopes present in the polypeptide; thus only a unique portion of a polypeptide (having the epitope) may need to be present in an assay to determine if the antibody specifically binds the polypeptide.

The antibody can be a component of a composition that comprises an antibody that specifically binds the AAV4 protein. The composition can further comprise, *e.g.*, serum, serum-free medium, or a pharmaceutically acceptable carrier such as physiological saline, etc..

By "an antibody that specifically binds" an AAV4 polypeptide or protein is meant an antibody that selectively binds to an epitope on any portion of the AAV4 peptide such that the antibody selectively binds to the AAV4 polypeptide, *i.e.*, such that the antibody binds specifically to the corresponding AAV4 polypeptide without significant background. Specific binding by an antibody further means that the antibody can be used to selectively remove the target polypeptide from a sample comprising the polypeptide or and can readily be determined by radioimmuno assay (RIA), bioassay, or enzyme-linked immunosorbant (ELISA) technology. An ELISA method effective for the detection of the specific antibody-antigen binding can, for example, be as follows:

(1) bind the antibody to a substrate; (2) contact the bound antibody with a sample containing the antigen; (3) contact the above with a secondary antibody bound to a detectable moiety (e.g., horseradish peroxidase enzyme or alkaline phosphatase enzyme); (4) contact the above with the substrate for the enzyme; (5) contact the above with a color reagent; (6) observe the color change.

An antibody can include antibody fragments such as Fab fragments which retain the binding activity. Antibodies can be made as described in, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1988). Briefly, purified antigen can be injected into an animal in an amount and in intervals sufficient to elicit an immune response. Antibodies can either be purified directly, or spleen cells can be obtained from the animal. The cells are then fused with an immortal cell line and screened for antibody secretion. Individual hybridomas are then propagated as individual clones serving as a source for a particular monoclonal antibody.

The present invention additionally provides a method of screening a cell for infectivity by AAV4 comprising contacting the cell with AAV4 and detecting the presence of AAV4 in the cells. AAV4 particles can be detected using any standard physical or biochemical methods. For example, physical methods that can be used for this detection include 1) polymerase chain reaction (PCR) for viral DNA or RNA, 2) direct hybridization with labeled probes, 3) antibody directed against the viral structural or non- structural proteins. Catalytic methods of viral detection include, but are not limited to, detection of site and strand specific DNA nicking activity of Rep proteins or replication of an AAV origin- containing substrate. Additional detection methods are outlined in Fields, *Virology*, Raven Press, New York, New York. 1996.

For screening a cell for infectivity by AAV4 wherein the presence of AAV4 in the cells is determined by nucleic acid hybridization methods, a nucleic acid probe for such detection can comprise, for example, a unique fragment of any of the AAV4 nucleic acids provided herein. The uniqueness of any nucleic acid probe can readily be

determined as described herein for unique nucleic acids. The nucleic acid can be, for example, the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO: 1, 3, 5, 6, 7, 12, 13, 14, 15, 17 or 19, or a unique fragment thereof.

The present invention includes a method of determining the suitability of an AAV4 vector for administration to a subject comprising administering to an antibody-containing sample from the subject an antigenic fragment of an isolated AAV4 capsid protein, and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the AAV4 vector to be unsuitable for use in the subject. The AAV4 capsid protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:4. An immunogenic fragment of an isolated AAV4 capsid protein can also be used in these methods. The AAV4 capsid protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:17. The AAV4 capsid protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:19.

Alternatively, or additionally; an antigenic fragment of an isolated AAV4 Rep protein can be utilized in this determination method. An immunogenic fragment of an isolated AAV4 Rep protein can also be used in these methods. Thus the present invention further provides a method of determining the suitability of an AAV4 vector for administration to a subject comprising administering to an antibody-containing sample from the subject an antigenic fragment of an AAV4 Rep protein and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the AAV4 vector to be unsuitable for use in the subject. The AAV4 Rep protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:2. The AAV4 Rep protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:8. The AAV4 Rep protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:9. The AAV4 Rep protein from which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:10. The AAV4 Rep protein from

which an antigenic fragment is selected can have the amino acid sequence set forth in SEQ ID NO:11.

An antigenic or immunoreactive fragment is typically an amino acid sequence of at least about 5 consecutive amino acids, and it can be derived from the AAV4 polypeptide amino acid sequence. An antigenic fragment is any fragment unique to the AAV4 protein, as described herein, against which an AAV4-specific antibody can be raised, by standard methods. Thus, the resulting antibody-antigen reaction should be specific for AAV4.

The AAV4 polypeptide fragments can be analyzed to determine their antigenicity, immunogenicity and/or specificity. Briefly, various concentrations of a putative immunogenically specific fragment are prepared and administered to a subject and the immunological response (e.g., the production of antibodies or cell mediated immunity) of an animal to each concentration is determined. The amounts of antigen administered depend on the subject, e.g. a human, rabbit or a guinea pig, the condition of the subject, the size of the subject, etc. Thereafter an animal so inoculated with the antigen can be exposed to the AAV4 viral particle or AAV4 protein to test the immunoreactivity or the antigenicity of the specific immunogenic fragment. The specificity of a putative antigenic or immunogenic fragment can be ascertained by testing sera, other fluids or lymphocytes from the inoculated animal for cross reactivity with other closely related viruses, such as AAV1, AAV2, AAV3 and AAV5.

As will be recognized by those skilled in the art, numerous types of immunoassays are available for use in the present invention to detect binding between an antibody and an AAV4 polypeptide of this invention. For instance, direct and indirect binding assays, competitive assays, sandwich assays, and the like, as are generally described in, *e.g.*, U.S. Pat. Nos. 4,642,285; 4,376,110; 4,016,043; 3,879,262; 3,852,157; 3,850,752; 3,839,153; 3,791,932; and Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, N.Y. (1988). For example, enzyme immunoassays such as immunofluorescence assays (IFA), enzyme linked

immunosorbent assays (ELISA) and immunoblotting can be readily adapted to accomplish the detection of the antibody. An ELISA method effective for the detection of the antibody bound to the antigen can, for example, be as follows: (1) bind the antigen to a substrate; (2) contact the bound antigen with a fluid or tissue sample containing the antibody; (3) contact the above with a secondary antibody specific for the antigen and bound to a detectable moiety (e.g., horseradish peroxidase enzyme or alkaline phosphatase enzyme); (4) contact the above with the substrate for the enzyme; (5) contact the above with a color reagent; (6) observe color change.

The antibody-containing sample of this method can comprise any biological sample which would contain the antibody or a cell containing the antibody, such as blood, plasma, serum, bone marrow, saliva and urine.

By the "suitability of an AAV4 vector for administration to a subject" is meant a determination of whether the AAV4 vector will elicit a neutralizing immune response upon administration to a particular subject. A vector that does not elicit a significant immune response is a potentially suitable vector, whereas a vector that elicits a significant, neutralizing immune response is thus indicated to be unsuitable for use in that subject. Significance of any detectable immune response is a standard parameter understood by the skilled artisan in the field. For example, one can incubate the subject's serum with the virus, then determine whether that virus retains its ability to transduce cells in culture. If such virus cannot transduce cells in culture, the vector likely has elicited a significant immune response.

The present method further provides a method of delivering a nucleic acid to a cell comprising administering to the cell an AAV4 particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell. Administration to the cell can be accomplished by any means, including simply contacting the particle, optionally contained in a desired liquid such as tissue culture medium, or a buffered saline solution, with the cells. The particle can be allowed to remain in contact with the cells for any

desired length of time, and typically the particle is administered and allowed to remain indefinitely. For such *in vitro* methods, the virus can be administered to the cell by standard viral transduction methods, as known in the art and as exemplified herein. Titers of virus to administer can vary, particularly depending upon the cell type, but will be typical of that used for AAV transduction in general. Additionally the titers used to transduce the particular cells in the present examples can be utilized. The cells can include any desired cell, such as the following cells and cells derived from the following tissues, in humans as well as other mammals, such as primates, horse, sheep, goat, pig, dog, rat, and mouse: Adipocytes, Adenocyte, Adrenal cortex, Amnion, Aorta, Ascites, Astrocyte, Bladder, Bone, Bone marrow, Brain, Breast, Bronchus, Cardiac muscle, Cecum, Cervix, Chorion, Colon, Conjunctiva, Connective tissue, Cornea, Dermis, Duodenum, Endometrium, Endothelium, Epithelial tissue, Epidermis, Esophagus, Eye, Fascia, Fibroblasts, Foreskin, Gastric, Glial cells, Glioblast, Gonad, Hepatic cells, Histocyte, Ileum, Intestine, small Intestine, Jejunum, Keratinocytes, Kidney, Larynx, Leukocytes, Lipocyte, Liver, Lung, Lymph node, Lymphoblast, Lymphocytes, Macrophages, Mammary alveolar nodule, Mammary gland, Mastocyte, Maxilla, Melanocytes, Monocytes, Mouth, Myelin, Nervous tissue, Neuroblast, Neurons, Neuroglia, Osteoblasts, Osteogenic cells, Ovary, Palate, Pancreas, Papilloma, Peritoneum, Pituicytes, Pharynx, Placenta, Plasma cells, Pleura, Prostate, Rectum, Salivary gland, Skeletal muscle, Skin, Smooth muscle, Somatic, Spleen, Squamous, Stomach, Submandibular gland, Submaxillary gland, Synoviocytes, Testis, Thymus, Thyroid, Trabeculae, Trachea, Turbinate, Umbilical cord, Ureter, and Uterus.

The AAV inverted terminal repeats in the vector for the herein described delivery methods can be AAV4 inverted terminal repeats. Specifically, they can comprise the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO:6 or the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO:20, or any fragment thereof demonstrated to have ITR functioning. The ITRs can also consist essentially of the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO:6 or the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO:20. Furthermore, the AAV inverted terminal repeats in the vector for the herein described nucleic acid delivery

methods can also comprise AAV2 inverted terminal repeats. Additionally, the AAV inverted terminal repeats in the vector for this delivery method can also consist essentially of AAV2 inverted terminal repeats.

The present invention also includes a method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject. The AAV ITRs can be any AAV ITRs, including AAV4 ITRs and AAV2 ITRs. For such an *ex vivo* administration, cells are isolated from a subject by standard means according to the cell type and placed in appropriate culture medium, again according to cell type (*see, e.g.*, ATCC catalog). Viral particles are then contacted with the cells as described above, and the virus is allowed to transfect the cells. Cells can then be transplanted back into the subject's body, again by means standard for the cell type and tissue (*e. g.*, in general, U.S. Patent No. 5,399,346; for neural cells, Dunnett, S.B. and Björklund, A., eds., *Transplantation: Neural Transplantation-A Practical Approach,* Oxford University Press, Oxford (1992)). If desired, prior to transplantation, the cells can be studied for degree of transfection by the virus, by known detection means and as described herein. Cells for *ex vivo* transfection followed by transplantation into a subject can be selected from those listed above, or can be any other selected cell. Preferably, a selected cell type is examined for its capability to be transfected by AAV4. Preferably, the selected cell will be a cell readily transduced with AAV4 particles; however, depending upon the application, even cells with relatively low transduction efficiencies can be useful, particularly if the cell is from a tissue or organ in which even production of a small amount of the protein or antisense RNA encoded by the vector will be beneficial to the subject.

The present invention further provides a method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject. Administration can be an *ex vivo*

administration directly to a cell removed from a subject, such as any of the cells listed above, followed by replacement of the cell back into the subject, or administration can be *in vivo* administration to a cell in the subject. For *ex vivo* administration, cells are isolated from a subject by standard means according to the cell type and placed in appropriate culture medium, again according to cell type (*see, e.g.*, ATCC catalog). Viral particles are then contacted with the cells as described above, and the virus is allowed to transfect the cells. Cells can then be transplanted back into the subject's body, again by means standard for the cell type and tissue (*e. g.*, for neural cells, Dunnett, S.B. and Björklund, A., eds., *Transplantation: Neural Transplantation-A Practical Approach*, Oxford University Press, Oxford (1992)). If desired, prior to transplantation, the cells can be studied for degree of transfection by the virus, by known detection means and as described herein.

In vivo administration to a human subject or an animal model can be by any of many standard means for administering viruses, depending upon the target organ, tissue or cell. Virus particles can be administered orally, parenterally (*e.g.*, intravenously), by intramuscular injection, by direct tissue or organ injection, by intraperitoneal injection, topically, transdermally, or the like. Viral nucleic acids (non-encapsidated) can be administered, *e.g.*, as a complex with cationic liposomes, or encapsulated in anionic liposomes. Compositions can include various amounts of the selected viral particle or non-encapsidated viral nucleic acid in combination with a pharmaceutically acceptable carrier and, in addition, if desired, may include other medicinal agents, pharmaceutical agents, carriers, adjuvants, diluents, etc. Parental administration, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as emulsions. Dosages will depend upon the mode of administration, the disease or condition to be treated, and the individual subject's condition, but will be that dosage typical for and used in administration of other AAV vectors, such as AAV2 vectors. Often a single dose can be sufficient; however, the dose can be repeated if desirable.

The present invention further provides a method of delivering a nucleic acid to a cell in a subject having antibodies to AAV2 comprising administering to the subject an AAV4 particle comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject. A subject that has antibodies to AAV2 can readily be determined by any of several known means, such as contacting AAV2 protein(s) with an antibody-containing sample, such as blood, from a subject and detecting an antigen-antibody reaction in the sample. Delivery of the AAV4 particle can be by either *ex vivo* or *in vivo* administration as herein described. Thus, a subject who might have an adverse immunogenic reaction to a vector administered in an AAV2 viral particle can have a desired nucleic acid delivered using an AAV4 particle. This delivery system can be particularly useful for subjects who have received therapy utilizing AAV2 particles in the past and have developed antibodies to AAV2. An AAV4 regimen can now be substituted to deliver the desired nucleic acid.

STATEMENT OF UTILITY

The present invention provides recombinant vectors based on AAV4. Such vectors may be useful for transducing erythroid progenitor cells which is very inefficient with AAV2 based vectors. In addition to transduction of other cell types, transduction of erythroid cells would be useful for the treatment of cancer and genetic diseases which can be corrected by bone marrow transplants using matched donors. Some examples of this type of treatment include, but are not limited to, the introduction of a therapeutic gene such as genes encoding interferons, interleukins, tumor necrosis factors, adenosine deaminase, cellular growth factors such as lymphokines, blood coagulation factors such as factor VIII and IX, cholesterol metabolism uptake and transport protein such as EpoE and LDL receptor, and antisense sequences to inhibit viral replication of, for example, hepatitis or HIV.

The present invention provides a vector comprising the AAV4 virus as well as AAV4 viral particles. While AAV4 is similar to AAV2, the two viruses are found herein to be physically and genetically distinct. These differences endow AAV4 with some unique advantages which better suit it as a vector for gene therapy. For example, the wt AAV4 genome is larger than AAV2, allowing for efficient encapsidation of a larger recombinant genome. Furthermore, wt AAV4 particles have a greater buoyant density than AAV2 particles and therefore are more easily separated from contaminating helper virus and empty AAV particles than AAV2-based particles.

Furthermore, as shown herein, AAV4 capsid protein is distinct from AAV2 capsid protein and exhibits different tissue tropism. AAV2 and AAV4 are shown herein to utilize distinct cellular receptors. AAV2 and AAV4 have been shown to be serologically distinct and thus, in a gene therapy application, AAV4 would allow for transduction of a patient who already possess neutralizing antibodies to AAV2 either as a result of natural immunological defense or from prior exposure to AAV2 vectors.

The present invention is more particularly described in the following examples which are intended as illustrative only since numerous modifications and variations therein will be apparent to those skilled in the art.

EXAMPLES

To understand the nature of AAV4 virus and to determine its usefulness as a vector for gene transfer, it was cloned and sequenced.

Cell culture and virus propagation

Cos and HeLa cells were maintained as monolayer cultures in D10 medium (Dulbecco's modified Eagle's medium containing 10% fetal calf serum, loo ug/ml penicillin, 100 units/ml streptomycin and IX Fungizone as recommended by the manufacturer; (GIBCO, Gaithersburg, MD, USA) . All other cell types were grown under standard conditions which have been previously reported. AAV4 stocks were obtained from American Type Culture Collection # VR- 64 6.

Virus was produced as previously described for AAV2 using the Beta galactosidase vector plasmid and a helper plasmid containing the AAV4 Rep and Cap genes (9). The helper plasmid was constructed in such a way as not to allow any homologous sequence between the helper and vector plasmids. This step was taken to minimize the potential for wild-type (wt) particle formation by homologous recombination.

Virus was isolated from $5x10^7$ cos cells by CsCl banding (9), and the distribution of Beta galactosidase genomes across the genome was determined by DNA dot blots of aliquots of gradient fractions. The majority of packaged genomes were found in fractions with a density of 1.43 which is similar to that reported for wt AAV4. This preparation of virus yielded 2.5 $X10^{11}$ particles or 5000 particles/producer cell. In comparison AAV2 isolated and CsCl banded from $8X10^7$ cells yielded 1.2 $X10^{11}$ particles or 1500 particles/producer cell. Thus, typical yields of rAAV4 particles/producer cell were 3-5 fold greater than that of rAAV2 particles.

DNA Cloning and Sequencing and Analysis

In order to clone the genome of AAV4, viral lysate was amplified in cos cells and then HeLa cells with the resulting viral particles isolated by CsCl banding. DNA dot blots of aliquots of the gradient fractions indicated that peak genomes were contained in fractions with a density of 1.41-1.45. This is very similar to the buoyant density previously reported for AAV4 (29). Analysis of annealed DNA obtained from these fractions indicated a major species of 4.8kb in length which upon restriction analysis gave bands similar in size to those previously reported. Additional restriction analysis indicated the presence of BssHII restriction sites near the ends of the DNA. Digestion with BssHII yielded a 4.5kb fragment which was then cloned into Bluescript SKII+ and two independent clones were sequenced.

The viral sequence is now available through Genebank, accession number U89790. DNA sequence was determined using an ABI 373A automated sequencer and the FS dye terminator chemistry. Both strands of the plasmids were sequenced and confirmed by sequencing of a second clone. As further confirmation of the authenticity of the sequence, bases 91-600 were PCR amlified from the original seed material and directly sequenced. The sequence of this region, which contains a 56 base insertion compared to AAV2 and 3, was found to be identical to that derived from the cloned material. The ITR was cloned using Deep Vent Polymerase (New England Biolabs) according to the manufactures instructions using the following primers, primer 1: 5'TCTAGTCTAGACTTGGCCACTCCCTCTCTGCGCGC(SEQ ID NO:21); primer 2: 51 AGGCCTTAAGAGCAGTCGTCCACCACCTTGTTCC (SEQ ID NO:22). Cycling conditions were 97°C 20 sec, 65°C 30 sec, 75°C 1 min for 35 rounds. Following the PCR reaction, the mixture was treated with XbaI and EcoRI endonucleases and the amplified band purified by agarose gel electrophoresis. The recovered DNA fragment was ligated into Bluescript SKII+ (Stratagene) and transformed into competent Sure strain bacteria (Stratagene). The helper plasmid ($pSV40oriAAV_{4-2}$) used for the production of recombinant virus, which contains the rep and cap genes of AAV4, was produced by PCR with *Pfu* polymerase (Stratagene)

according to the manufactures instructions. The amplified sequence, nt 216-4440, was ligated into a plasmid that contains the SV40 origin of replication previously described (9, 10). Cycling conditions were 95°C 30 sec, 55°C 30 sec, 72°C 3 min for 20 rounds. The final clone was confirmed by sequencing. The βgal reporter vector has been described previously (9, 10).

Sequencing of this fragment revealed two open reading frames (ORF) instead of only one as previously suggested. In addition to the previously identified Capsid ORF in the right-hand side of the genome, an additional ORF is present on the left-hand side. Computer analysis indicated that the left-hand ORF has a high degree of homology to the Rep gene of AAV2. At the amino acid level the ORF is 90% identical to that of AAV2 with only 5% of the changes being non-conserved (SEQ ID NO:2). In contrast, the right ORF is only 62% identical at the amino acid level when compared to the corrected AAV2 sequence. While the internal start site of VP2 appears to be conserved, the start site for VP3 is in the middle of one of the two blocks of divergent sequence. The second divergent block is in the middle of VP3. By using three dimensional structure analysis of the canine parvovirus and computer aided sequence comparisons, regions of AAV2 which might be exposed on the surface of the virus have been identified. Comparison of the AAV2 and AAV4 sequences indicates that these regions are not well conserved between the two viruses and suggests altered tissue tropism for the two viruses.

Comparison of the p5 promoter region of the two viruses shows a high degree of conservation of known functional elements (SEQ ID NO:7). Initial work by Chang *et al.* identified two YY1 binding sites at -60 and +1 and a TATA Box at -30 which are all conserved between AAV2 and AAV4 (4). A binding site for the Rep has been identified in the p5 promoter at -17 and is also conserved (24). The only divergence between the two viruses in this region appears to be in the sequence surrounding these elements. AAV4 also contains an additional 56 bases in this region between the p5 promoter and the TRS (nt 209-269). Based on its positioning in the viral genome and efficient use of the limited genome space, this sequence may possess some promoter activity or be involved in rescue, replication or packaging of the virus.

The inverted terminal repeats were cloned by PCR using a probe derived from the terminal resolution site (TRS)of the BssHII fragment and a primer in the Rep ORF. The TRS is a sequence at the end of the stem of the ITR and the reverse compliment of TRS sequence was contained within the BssHII fragment. The resulting fragments were cloned and found to contain a number of sequence changes compared to AAV2. However, these changes were found to be complementary and did not affect the ability of this region to fold into a hairpin structure (Fig 2). While the TRS site was conserved between AAV2 and AAV4 the Rep binding site contained two alterations which expand the binding site from 3 GAGC repeats to 4. The first two repeats in AAV4 both contain a T in the fourth position instead of a C. This type of repeat is present in the p5 promoter and is present in the consensus sequence that has been proposed for Rep binding (10) and its expansion may affect its affinity for Rep. Methylation interference data has suggested the importance of the CTTTG motif found at the tip of one palindrome in Rep binding with the underlined T residues clearly affecting Rep binding to both the flip and flop forms. While most of this motif is conserved in AAV4 the middle T residue is changed to a C (33).

Hemagglutination assays

Hemagglutination was measured essentially as described previously (18). Serial two fold dilutions of virus in Veronal-buffered saline were mixed with an equal volume of 0.4% human erythrocytes (type 0) in plastic U bottom 96 well plates. The reaction was complete after a 2 hr incubation at 8°C. HA units (HAU) are defined as the reciprocal of the dilution causing 50% hemagglutination.

The results show that both the wild type and recombinant AAV4 viruses can hemagglutinate human red blood cells (RBCS) with HA titers of approximately 1024 HAU/μl and 512 HAU/μl respectively. No HA activity was detected with AAV type 3 or recombinant AAV type 2 as well as the helper adenovirus. If the temperature was raised to 22°C, HA activity decreased 32-fold. Comparison of the viral particle number per RBC at the end point dilution indicated that approximately 1-10 particles per RBC were required for hemagglutination. This value is similar to that previously reported (18).

Tissue tropism analysis

The sequence divergence in the capsid proteins ORF which are predicted to be exposed on the surface of the virus may result in an altered binding specificity for AAV4 compared to AAV2. Very little is known about the tissue tropism of any dependovirus. While it had been shown to hemagglutinate human, guinea pig, and sheep erythrocytes, it is thought to be exclusively a simian virus (18). Therefore, to examine AAV4 tissue tropism and its species specificity, recombinant AAV4 particles which contained the gene for nuclear localized Beta galactosidase were constructed. Because of the similarity in genetic organization of AAV4 and AAV2, it was determined whether AAV4 particles could be constructed containing a recombinant genome. Furthermore, because of the structural similarities of the AAV type 2 and type 4 ITRs, a genome containing AAV2 ITRs which had been previously described was used.

Tissue tropism analysis 1. To study AAV transduction, a variety of cell lines were transduced with 5 fold serial dilutions of either recombinant AAV2 or AAV4 particles expressing the gene for nuclear localized Beta galactosidase activity (Table 1). Approximately 4 $X10^4$ cells were exposed to virus in 0.5ml serum free media for 1 hour and then 1 ml of the appropriate complete media was added and the cells were incubated for 48-60 hours. The cells were then fixed and stained for β-galactosidase activity with 5-Bromo-4-Chloro-3-Indolyl-β-D-galactopyranoside (Xgal) (ICN Biomedicals) (36). Biological titers were determined by counting the number of positive cells in the different dilutions using a calibrated microscope ocular ($3.1mm^2$) then multiplying by the area of the well and the dilution of the virus. Typically dilutions which gave 1-10 positive cells per field (100-1000 positive cells per 2cm well) were used for titer determination. Titers were determined by the average number of cells in a minimum of 10 fields/well.

To examine difference in tissue tropism, a number of cell lines were transduced with serial dilutions of either AAV4 or AAV2 and the biological titers determined. As shown in Table 1, when Cos cells were transduced with a similar number of viral particles, a similar level of transduction was observed with AAV2 and AAV4.

However, other cell lines exhibited differential transducibility by AAV2 or AAV4. Transduction of the human colon adenocarcinoma cell line SW480 with AAV2 was over 100 times higher than that obtained with AAV4. Furthermore, both vectors transduced SW1116, SW1463 and NIH3T3 cells relatively poorly.

TABLE 1

Cell type	AAV2	AAV4
Cos	4.5 X10^7	1.9 X10^7
SW 480	3.8 X10^6	2.8 X10^4
SW 1116	5.2 X10^4	8 X10^3
SW1463	8.8 X10^4	8 X10^3
SW620	8.8 X10^4	ND
NIH 3T3	2 X10^4	8X10^3

Tissue tropism analysis 2.

A. Transduction of cells. Exponentially growing cells (2 X 10^4) were plated in each well of a 12 well plate and transduced with serial dilutions of virus in 200 µl of medium for I hr. After this period, 800 µl of additional medium was added and incubated for 48 hrs. The cells were then fixed and stained for β-galactosidase activity overnight with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal) (ICN Biomedicals) (36). No endogenous β-galactosidase activity was visible after 24 hr incubation in Xgal solution. Infectious titers were determined by counting the number of positive cells in the different dilutions using a calibrated microscope ocular (diameter 3.1 mm^2) then multiplying by the area of the well and the dilution of the virus. Titers were determined by the average number of cells in a minimum of 10 fields/well.

As shown in Table 2, cos cells transduced with equivalent amounts of rAAV2 and rAAV4particles resulted in similar transduction levels. However, other cell lines exhibited differential transducibility. Transduction of the human colon adenocarcinoma cell line, SW480, with rAAV2 was 60 times higher than that obtained with rAAV4. Hela

and SW620 cells were also transduced more efficiently with rAAV2 than rAAV4. In contrast, transduction of primary rat brain cultures exhibited a greater transduction of glial and neuronal cells with rAAV4 compared to rAAV2. Because of the heterogeneous nature of the cell population in the rat brain cultures, only relative transduction efficiencies are reported (Table 2).

As a control for adenovirus contamination of the viral preparations cos and Hela cells were coinfected with RAAV and adenovirus then stained after 24 hr. While the titer of rAAV2 increased in the presence of Ad in both cos and Hela, adenovirus only increased the titer in the cos cells transduced with rAAV4 and not the HeLa cells, suggesting the difference in transduction efficiencies is not the result of adenovirus contamination. Furthermore, both vectors transduced SW1116, SW1463, NIH3T3 and monkey fibroblasts FL2 cells very poorly. Thus AAV4 may utilize a cellular receptor distinct from that of AAV2.

TABLE 2

CELL TYPE	AAV2	AAV4
Primary Rat Brain	1	4.3 ± 0.7
cos	$4.2X10^7\pm4.6X10^6$	$2.2X10^7\pm2.5X10^6$
SW 480	$7.75X10^6\pm1.7X10^6$	$1.3X10^5\pm6.8X10^4$
Hela	$2.1X10^7\pm1X10^6$	$1.3X10^6\pm1X10^5$
SW620	$1.2X10^5\pm3.9X10^4$	$4X10^4$
KLEB	$1.2X10^5\pm3.5X10^4$	$9X10^4\pm1.4X10^4$
HB	$5.6X10^5\pm2X10^5$	$3.8X10^4\pm1.8X10^4$
SW1116	$5.2 X 10^4$	$8 X 10^3$
SW1463	$8.8 X 10^4$	$8 X 10^3$
NIH 3T3	$3 X 10^3$	$2 X 10^3$

B. Competition assay. Cos cells were plated at 2x 10^4 /well in 12 well plates 12-24 hrs prior to transduction. Cells were transduced with 0.5x 10^7 particles of rAAV2 or rAAV4 (containing the LacZ gene) in 200 µl of DMEM and increasing amounts of rAAV2 containing the gene for the human coagulation factor IX. Prior to transduction the CsCl was removed from the virus by dialysis against isotonic saline. After lhr incubation with the recombinant virus the culture medium was supplemented with complete medium and allowed to incubate for 48-60 hrs. The cells were then stained and counted as described above.

AAV4 utilization of a cellular receptor distinct from that of AAV2 was further examined by cotransduction experiments with rAAV2 and rAAV4. Cos cells were transduced with an equal number of rAAV2 or rAAV4 particles containing the LacZ gene and increasing amounts of rAAV2 particles containing the human coagulation factor IX gene (rAAV2FIX) . At a 72:1 ratio of rAAV2FIX:rAAV4LacZ only a two-fold effect on the level of rAAV4LacZ transduction was obtained (Fig 3). However this same ratio of rAAV2FIX:rAAV2LacZ reduced the transduction efficiency of rAAV2LacZ approximately 10 fold. Comparison of the 50% inhibition points for the two viruses indicated a 7 fold difference in sensitivity.

C. Trypsinization of cells. An 80% confluent monolayer of cos cells (1x 10^7) was treated with 0.05% trypsin/0.02% versene solution (Biofluids) for 3-5 min at 37°C. Following detachment the trypsin was inactivated by the addition of an equal volume of media containing 10% fetal calf serum. The cells were then further diluted to a final concentration of 1x 10^4/ml. One ml of cells was plated in a 12 well dish and incubated with virus at a multiplicity of infection (MOI) of 260 for 1-2 hrs. Following attachment of the cells the media containing the virus was removed, the cells washed and fresh media was added. Control cells were plated at the same time but were not transduced until the next day. Transduction conditions were done as described above for the trypsinized cell group. The number of transduced cells was determined by staining 48-60 hrs post transduction and counted as described above.

Previous research had shown that binding and infection of AAV2 is inhibited by trypsin treatment of cells (26). Transduction of cos cells with rAAV2lacZ gene was also inhibited by trypsin treatment prior to transduction (Fig 4). In contrast trypsin treatment had a minimal effect on rAAV4lacZ transduction. This result and the previous competition experiment are both consistent with the utilization of distinct cellular receptors for AAV2 and AAV4.

AAV4 is a distinct virus based on sequence analysis, physical properties of the virion, hemagglutination activity, and tissue tropism. The sequence data indicates that AAV4 is a distinct virus from that of AAV2. In contrast to original reports, AAV4 contains two open reading frames which code for either Rep proteins or Capsid proteins. AAV4 contains additional sequence upstream of the p5 promoter which may affect promoter activity, packaging or particle stability. Furthermore, AAV4 contains an expanded Rep binding site in its ITR which could alter its activity as an origin of replication or promoter. The majority of the differences in the Capsid proteins lies in regions which have been proposed to be on the exterior surface of the parvovirus. These changes are most likely responsible for the lack of cross reacting antibodies, hemagglutinate activity, and the altered tissue tropism compared to AAV2. Furthermore, in contrast to previous reports AAV4 is able to transduce human as well as monkey cells.

Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

Although the present process has been described with reference to specific details of certain embodiments thereof, it is not intended that such details should be regarded as limitations upon the scope of the invention except as and to the extent that they are included in the accompanying claims.

References:

1. Arella, M., S. Garzon, J. Bergeron, and P. Tijssen. Handbook of Parvoviruses. Vol. 1. ed. P. Tijssen. Boca Raton, Florida, CRC Press, 1990.
2. Bachmann, P.A., M.D. Hoggan, E. Kurstak, J.L. Melnick, H.G. Pereira, P. Tattersall, and C. Vago. 1979. Interverology 11: 248-254.
3. Bantel-Schaal, U. and M. Stohr. 1992. J. Virol. 66: 773-779.
4. Chang, L.S., Y. Shi, and T. Shenk. 1989. J. Virol. 63: 3479-88.
5. Chejanovsky, N. and B.J. Carter. 1989. Virology 173: 120-128.
6. Chejanovsky, N. and B.J. Carter. 1989. Virology 171: 239-247.
7. Chiorini, J.A., S.M. Wiener, R.M. Kotin, R.A. Owens, SRM Kyöstiö, and B. Safer. 1994. J. Virol. 68: 7448-7457.
8. Chiorini, J.A., M.D. Weitzman, R.A. Owens, E. Urcelay, B. Safer, and R.M. Kotin. 1994. J. Virol. 68: 797-804.
9. Chiorini, J.A., C.M. Wendtner, E. Urcelay, B. Safer, M. Hallek, and R.M. Kotin. 1995. Human Gene Therapy 6: 1531-1541.
10. Chiorini, J.A., L. Yang, B. Safer, and R.M. Kotin. 1995. J. Virol. 69: 7334-7338.
11. Dixit, M., M.S. Webb, W.C. Smart, and S. Ohi. 1991. Gene 104: 253-7.
12. Fisher, R.E. and H.D. Mayor. 1991. J Theor Biol 149: 429-39.
13. Flotte, T.R., S.A. Afione, C. Conrad, S.A. McGrath, R. Solow, H. Oka, P.L. Zeitlin, W.B. Guggino, and B.J. Carter. 1993. Proc. Natl. Acad. Sci. 90: 10613-10617.
14. Flotte, T.R., S.A. Afione, R. Solow, M.L. Drumm, D. Markakis, W.B. Guggino, P.L. Zeitlin, and B.J. Carter. 1993. J Biol Chem 268: 3781-90.
15. Hermonat, P.L., M.A. Labow, R. Wright, K.I. Berns, and N. Muzyczka. 1984. J. Virol. 51: 329-339.

16. **Hermonat, P.L. and N. Muzyczka.** 1984. Proc Natl Acad Sci USA **81:** 6466-70.

17. **Hunter, L.A. and R.J. Samulski.** 1992. J. Virol. **66:** 317-24.

18. **Ito, M. and H.D. Mayor.** 1968. J. Immuno. **100:** 61-68.

19. **Janik, J.E., M.M. Huston, K. Cho, and J.A. Rose.** 1989. Virology **168:** 320-9.

20. **Kaplitt, M.G., P. Leone, R.J. Samulski, X. Xiao, D.W. Pfaff, K.L. O'Malley, and J.M. During.** 1994. Nature Genetics **8:** 148-154.

21. **Kotin, R.M., M. Siniscalco, R.J. Samulski, X. Zhu, L. Hunter, C.A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K.I. Berns.** 1990. Proc. Natl. Acad. Sci. (USA) **87:** 2211-2215.

22. **Laughlin, C.A., N. Jones, and B.J. Carter.** 1982. J. Virol. **41:** 868-76.

23. **Laughlin, C.A., M.W. Myers, D.L. Risin, B.J. Carter.** 1979. Virology **94:** 162-74.

24. **McCarty, D.M., J. Pereira, I. Zolotukhin, X. Zhou, J.H. Ryan, and N. Muzyczka.** 1994. J. Virol. **68:** 4988-4997.

25. **Mendelson, E., J.P. Trempe, and B.J. Carter.** 1986. J. Virol. **60:** 823-832.

26. **Mizukami, H., N.S. Young, and K.E. Brown.** 1996. Virology **217:** 124-130.

27. **Muster, C.J., Y.S. Lee, J.E. Newbold, and J. Leis.** 1980. J. Virol. **35:** 653-61.

28. **Muzyczka, N.** 1992. Curr Top Microbiol Immunol **158:** 97-129.

29. **Parks, W.P., J.L. Melnick, R. Rongey, and H.D. Mayor.** 1967. J. Virol. **1:** 171-180.

30. **Podsakoff, G., K.K. Jr Wong, and S. Chatterjee.** 1994. J. Virol. **68:** 5656-5666.

31. **Rose, J.A., M.D. Hoggan, F. Koczot, and A.J. Shatkin.** 1968. J. Virol. **2:** 999-1005.

32. **Russell, D.W., A.D. Miller, and I.E. Alexander.** 1994. Proc. Natl. Acad. Sci. USA **91:** 8915-8919.

33. **Ryan, J.H., S. Zolotukhin, and N. Muzyczka.** 1996. J. Virol. **70:** 1542-1553.

34. **Samulski, R.J., K.I. Berns, M. Tan, and N. Muzyczka.** 1982. Proc Natl Acad Sci USA **79:** 2077-81.

35. **Samulski, R.J., L.S. Chang, and T. Shenk.** 1989. J. Virol. **63:** 3822-8.

36. **Sanes, J.R., J.L.R. Rubenstein, and J.F. Nicocas.** 1986. EMBO **5:** 3133-3142.

37. **Senapathy, P., J.D. Tratschin, and B.J. Carter.** 1984. J Mol Biol **179:** 1-20.

38. **Tratschin, J.D., I.L. Miller, and B.J. Carter.** 1984. J. Virol. **51:** 611-619.

39. **Trempe, J.P. and B.J. Carter.** 1988. J. Virol. **62:** 68-74.

40. **Trempe, J.P., E. Mendelson, and B.J. Carter.** 1987. Virology **161:** 18-28.

41. **Walsh, C.E., J.M. Liu, X. Xiao, N.S. Young, A.W. Nienhuis, and R.J. Samulski.** 1992. Proc Natl Acad Sci USA **89:** 7257-61.

42. **Winocour, E., M.F. Callaham, and E. Huberman.** 1988. Virology **167:** 393-9.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Chiorini, John A.
Kotin, Robert M.
Safer, Brian

(ii) TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Needle & Rosenberg
(B) STREET: 127 Peachtree
(C) CITY: Atlanta
(D) STATE: Georgia
(E) COUNTRY: USA
(F) ZIP: 30303

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Selby, Elizabeth
(B) REGISTRATION NUMBER: 38,298
(C) REFERENCE/DOCKET NUMBER: 14014.0252

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4767 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) OTHER INFO: AAV4 genome

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
TTGGCCACTC CCTCTATGCG CGCTCGCTCA CTCACTCGGC CCTGGAGACC AAAGGTCTCC    60

AGACTGCCGG CCTCTGGCCG GCAGGGCCGA GTGAGTGAGC GAGCGCGCAT AGAGGGAGTG   120

GCCAACTCCA TCATCTAGGT TTGCCCACTG ACGTCAATGT GACGTCCTAG GGTTAGGGAG   180

GTCCCTGTAT TAGCAGTCAC GTGAGTGTCG TATTTCGCGG AGCGTAGCGG AGCGCATACC   240

AAGCTGCCAC GTCACAGCCA CGTGGTCCGT TTGCGACAGT TTGCGACACC ATGTGGTCAG   300

GAGGGTATAT AACCGCGAGT GAGCCAGCGA GGAGCTCCAT TTTGCCCGCG AATTTTGAAC   360
```

GAGCAGCAGC CATGCCGGGG TTCTACGAGA TCGTGCTGAA GGTGCCCAGC GACCTGGACG 420
AGCACCTGCC CGGCATTTCT GACTCTTTTG TGAGCTGGGT GGCCGAGAAG GAATGGGAGC 480
TGCCGCCGGA TTCTGACATG GACTTGAATC TGATTGAGCA GGCACCCCTG ACCGTGGCCG 540
AAAAGCTGCA ACGCGAGTTC CTGGTCGAGT GGCGCCGCGT GAGTAAGGCC CCGGAGGCCC 600
TCTTCTTTGT CCAGTTCGAG AAGGGGGACA GCTACTTCCA CCTGCACATC CTGGTGGAGA 660
CCGTGGGCGT CAAATCCATG GTGGTGGGCC GCTACGTGAG CCAGATTAAA GAGAAGCTGG 720
TGACCCGCAT CTACCGCGGG GTCGAGCCGC AGCTTCCGAA CTGGTTCGCG GTGACCAAGA 780
CGCGTAATGG CGCCGGAGGC GGGAACAAGG TGGTGGACGA CTGCTACATC CCCAACTACC 840
TGCTCCCCAA GACCCAGCCC GAGCTCCAGT GGGCGTGGAC TAACATGGAC CAGTATATAA 900
GCGCCTGTTT GAATCTCGCG GAGCGTAAAC GGCTGGTGGC GCAGCATCTG ACGCACGTGT 960
CGCAGACGCA GGAGCAGAAC AAGGAAAACC AGAACCCCAA TTCTGACGCG CCGGTCATCA 1020
GGTCAAAAAC CTCCGCCAGG TACATGGAGC TGGTCGGGTG GCTGGTGGAC CGCGGGATCA 1080
CGTCAGAAAA GCAATGGATC CAGGAGGACC AGGCGTCCTA CATCTCCTTC AACGCCGCCT 1140
CCAACTCGCG GTCACAAATC AAGGCCGCGC TGGACAATGC CTCCAAAATC ATGAGCCTGA 1200
CAAAGACGGC TCCGGACTAC CTGGTGGGCC AGAACCCGCC GGAGGACATT TCCAGCAACC 1260
GCATCTACCG AATCCTCGAG ATGAACGGGT ACGATCCGCA GTACGCGGCC TCCGTCTTCC 1320
TGGGCTGGGC GCAAAAGAAG TTCGGGAAGA GGAACACCAT CTGGCTCTTT GGGCCGGCCA 1380
CGACGGGTAA AACCAACATC GCGGAAGCCA TCGCCCACGC CGTGCCCTTC TACGGCTGCG 1440
TGAACTGGAC CAATGAGAAC TTTCCGTTCA ACGATTGCGT CGACAAGATG GTGATCTGGT 1500
GGGAGGAGGG CAAGATGACG GCCAAGGTCG TAGAGAGCGC CAAGGCCATC CTGGGCGGAA 1560
GCAAGGTGCG CGTGGACCAA AAGTGCAAGT CATCGGCCCA GATCGACCCA ACTCCCGTGA 1620
TCGTCACCTC CAACACCAAC ATGTGCGCGG TCATCGACGG AAACTCGACC ACCTTCGAGC 1680
ACCAACAACC ACTCCAGGAC CGGATGTTCA AGTTCGAGCT CACCAAGCGC CTGGAGCACG 1740
ACTTTGGCAA GGTCACCAAG CAGGAAGTCA AAGACTTTTT CCGGTGGGCG TCAGATCACG 1800
TGACCGAGGT GACTCACGAG TTTTACGTCA GAAAGGGTGG AGCTAGAAAG AGGCCCGCCC 1860
CCAATGACGC AGATATAAGT GAGCCCAAGC GGGCCTGTCC GTCAGTTGCG CAGCCATCGA 1920
CGTCAGACGC GGAAGCTCCG GTGGACTACG CGGACAGGTA CCAAAACAAA TGTTCTCGTC 1980
ACGTGGGTAT GAATCTGATG CTTTTTCCCT GCCGGCAATG CGAGAGAATG AATCAGAATG 2040
TGGACATTTG CTTCACGCAC GGGGTCATGG ACTGTGCCGA GTGCTTCCCC GTGTCAGAAT 2100
CTCAACCCGT GTCTGTCGTC AGAAAGCGGA CGTATCAGAA ACTGTGTCCG ATTCATCACA 2160
TCATGGGGAG GGCGCCCGAG GTGGCCTGCT CGGCCTGCGA ACTGGCCAAT GTGGACTTGG 2220
ATGACTGTGA CATGGAACAA TAAATGACTC AAACCAGATA TGACTGACGG TTACCTTCCA 2280

GATTGGCTAG AGGACAACCT CTCTGAAGGC GTTCGAGAGT GGTGGGCGCT GCAACCTGGA 2340

GCCCCTAAAC CCAAGGCAAA TCAACAACAT CAGGACAACG CTCGGGGTCT TGTGCTTCCG 2400

GGTTACAAAT ACCTCGGACC CGGCAACGGA CTCGACAAGG GGGAACCCGT CAACGCAGCG 2460

GACGCGGCAG CCCTCGAGCA CGACAAGGCC TACGACCAGC AGCTCAAGGC CGGTGACAAC 2520

CCCTACCTCA AGTACAACCA CGCCGACGCG GAGTTCCAGC AGCGGCTTCA GGGCGACACA 2580

CCGTTTGGGG GCAACCTCGG CAGAGCAGTC TTCCAGGCCA AAAAGAGGGT TCTTGAACCT 2640

CTTGGTCTGG TTGAGCAAGC GGGTGAGACG GCTCCTGGAA AGAAGAGACC GTTGATTGAA 2700

TCCCCCCAGC AGCCCGACTC CTCCACGGGT ATCGGCAAAA AAGGCAAGCA GCCGGCTAAA 2760

AAGAAGCTCG TTTTCGAAGA CGAAACTGGA GCAGGCGACG GACCCCCTGA GGGATCAACT 2820

TCCGGAGCCA TGTCTGATGA CAGTGAGATG CGTGCAGCAG CTGGCGGAGC TGCAGTCGAG 2880

GGSGGACAAG GTGCCGATGG AGTGGGTAAT GCCTCGGGTG ATTGGCATTG CGATTCCACC 2940

TGGTCTGAGG GCCACGTCAC GACCACCAGC ACCAGAACCT GGGTCTTGCC CACCTACAAC 3000

AACCACCTNT ACAAGCGACT CGGAGAGAGC CTGCAGTCCA ACACCTACAA CGGATTCTCC 3060

ACCCCCTGGG GATACTTTGA CTTCAACCGC TTCCACTGCC ACTTCTCACC ACGTGACTGG 3120

CAGCGACTCA TCAACAACAA CTGGGGCATG CGACCCAAAG CCATGCGGGT CAAAATCTTC 3180

AACATCCAGG TCAAGGAGGT CACGACGTCG AACGGCGAGA CAACGGTGGC TAATAACCTT 3240

ACCAGCACGG TTCAGATCTT TGCGGACTCG TCGTACGAAC TGCCGTACGT GATGGATGCG 3300

GGTCAAGAGG GCAGCCTGCC TCCTTTTCCC AACGACGTCT TTATGGTGCC CCAGTACGGC 3360

TACTGTGGAC TGGTGACCGG CAACACTTCG CAGCAACAGA CTGACAGAAA TGCCTTCTAC 3420

TGCCTGGAGT ACTTTCCTTC GCAGATGCTG CGGACTGGCA ACAACTTTGA AATTACGTAC 3480

AGTTTTGAGA AGGTGCCTTT CCACTCGATG TACGCGCACA GCCAGAGCCT GGACCGGCTG 3540

ATGAACCCTC TCATCGACCA GTACCTGTGG GGACTGCAAT CGACCACCAC CGGAACCACC 3600

CTGAATGCCG GGACTGCCAC CACCAACTTT ACCAAGCTGC GGCCTACCAA CTTTTCCAAC 3660

TTTAAAAAGA ACTGGCTGCC CGGGCCTTCA ATCAAGCAGC AGGGCTTCTC AAAGACTGCC 3720

AATCAAAACT ACAAGATCCC TGCCACCGGG TCAGACAGTC TCATCAAATA CGAGACGCAC 3780

AGCACTCTGG ACGGAAGATG GAGTGCCCTG ACCCCCGGAC CTCCAATGGC CACGGCTGGA 3840

CCTGCGGACA GCAAGTTCAG CAACAGCCAG CTCATCTTTG CGGGGCCTAA ACAGAACGGC 3900

AACACGGCCA CCGTACCCGG GACTCTGATC TTCACCTCTG AGGAGGAGCT GGCAGCCACC 3960

AACGCCACCG ATACGGACAT GTGGGGCAAC CTACCTGGCG GTGACCAGAG CAACAGCAAC 4020

CTGCCGACCG TGGACAGACT GACAGCCTTG GGAGCCGTGC CTGGAATGGT CTGGCAAAAC 4080

AGAGACATTT ACTACCAGGG TCCCATTTGG GCCAAGATTC CTCATACCGA TGGACACTTT 4140

CACCCCTCAC CGCTGATTGG TGGGTTTGGG CTGAAACACC CGCCTCCTCA AATTTTTATC 4200

AAGAACACCC CGGTACCTGC GAATCCTGCA ACGACCTTCA GCTCTACTCC GGTAAACTCC 4260

TTCATTACTC AGTACAGCAC TGGCCAGGTG TCGGTGCAGA TTGACTGGGA GATCCAGAAG 4320

GAGCGGTCCA AACGCTGGAA CCCCGAGGTC CAGTTTACCT CCAACTACGG ACAGCAAAAC 4380

TCTCTGTTGT GGGCTCCCGA TGCGGCTGGG AAATACACTG AGCCTAGGGC TATCGGTACC 4440

CGCTACCTCA CCCACCACCT GTAATAACCT GTTAATCAAT AAACCGGTTT ATTCGTTTCA 4500

GTTGAACTTT GGTCTCCGTG TCCTTCTTAT CTTATCTCGT TTCCATGGCT ACTGCGTACA 4560

TAAGCAGCGG CCTGCGGCGC TTGCGCTTCG CGGTTTACAA CTGCCGGTTA ATCAGTAACT 4620

TCTGGCAAAC CAGATGATGG AGTTGGCCAC ATTAGCTATG CGCGCTCGCT CACTCACTCG 4680

GCCCTGGAGA CCAAAGGTCT CCAGACTGCC GGCCTCTGGC CGGCAGGGCC GAGTGAGTGA 4740

GCGAGCGCGC ATAGAGGGAG TGGCCAA 4767

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 624 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) OTHER INFO: AAV4 Rep protein (full length)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Pro Gly Phe Tyr Glu Ile Val Leu Lys Val Pro Ser Asp Leu Asp
1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu
20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Glu Phe Leu
50 55 60

Val Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65 70 75 80

Gln Phe Glu Lys Gly Asp Ser Tyr Phe His Leu His Ile Leu Val Glu
85 90 95

Thr Val Gly Val Lys Ser Met Val Val Gly Arg Tyr Val Ser Gln Ile
100 105 110

Lys Glu Lys Leu Val Thr Arg Ile Tyr Arg Gly Val Glu Pro Gln Leu
115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130 135 140

Asn Lys Val Val Asp Asp Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145 150 155 160

```
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Asp Gln Tyr Ile
                165                 170                 175

Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
            180                 185                 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
        195                 200                 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
    210                 215                 220

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
225                 230                 235                 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
                245                 250                 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
            260                 265                 270

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
        275                 280                 285

Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
    290                 295                 300

Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305                 310                 315                 320

Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
                325                 330                 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
            340                 345                 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
        355                 360                 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
    370                 375                 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385                 390                 395                 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
                405                 410                 415

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
            420                 425                 430

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
        435                 440                 445

Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
    450                 455                 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465                 470                 475                 480

Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
                485                 490                 495
```

```
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
            500                 505                 510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
        515                 520                 525

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
    530                 535                 540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Val Asp Ile Cys
545                 550                 555                 560

Phe Thr His Gly Val Met Asp Cys Ala Glu Cys Phe Pro Val Ser Glu
                565                 570                 575

Ser Gln Pro Val Ser Val Val Arg Lys Arg Thr Tyr Gln Lys Leu Cys
            580                 585                 590

Pro Ile His His Ile Met Gly Arg Ala Pro Glu Val Ala Cys Ser Ala
        595                 600                 605

Cys Glu Leu Ala Asn Val Asp Leu Asp Asp Cys Asp Met Glu Gln  *
    610                 615                 620
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1872 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) OTHER INFO: AAV4 Rep gene (full length)

(ix) FEATURE:
- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1872

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATG CCG GGG TTC TAC GAG ATC GTG CTG AAG GTG CCC AGC GAC CTG GAC       48
Met Pro Gly Phe Tyr Glu Ile Val Leu Lys Val Pro Ser Asp Leu Asp
  1               5                  10                  15

GAG CAC CTG CCC GGC ATT TCT GAC TCT TTT GTG AGC TGG GTG GCC GAG       96
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu
             20                  25                  30

AAG GAA TGG GAG CTG CCG CCG GAT TCT GAC ATG GAC TTG AAT CTG ATT      144
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
         35                  40                  45

GAG CAG GCA CCC CTG ACC GTG GCC GAA AAG CTG CAA CGC GAG TTC CTG      192
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Glu Phe Leu
     50                  55                  60
```

```
GTC GAG TGG CGC CGC GTG AGT AAG GCC CCG GAG GCC CTC TTC TTT GTC      240
Val Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65                  70                  75                  80

CAG TTC GAG AAG GGG GAC AGC TAC TTC CAC CTG CAC ATC CTG GTG GAG      288
Gln Phe Glu Lys Gly Asp Ser Tyr Phe His Leu His Ile Leu Val Glu
                 85                  90                  95

ACC GTG GGC GTC AAA TCC ATG GTG GTG GGC CGC TAC GTG AGC CAG ATT      336
Thr Val Gly Val Lys Ser Met Val Val Gly Arg Tyr Val Ser Gln Ile
            100                 105                 110

AAA GAG AAG CTG GTG ACC CGC ATC TAC CGC GGG GTC GAG CCG CAG CTT      384
Lys Glu Lys Leu Val Thr Arg Ile Tyr Arg Gly Val Glu Pro Gln Leu
        115                 120                 125

CCG AAC TGG TTC GCG GTG ACC AAG ACG CGT AAT GGC GCC GGA GGC GGG      432
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
    130                 135                 140

AAC AAG GTG GTG GAC GAC TGC TAC ATC CCC AAC TAC CTG CTC CCC AAG      480
Asn Lys Val Val Asp Asp Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145                 150                 155                 160

ACC CAG CCC GAG CTC CAG TGG GCG TGG ACT AAC ATG GAC CAG TAT ATA      528
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Asp Gln Tyr Ile
                165                 170                 175

AGC GCC TGT TTG AAT CTC GCG GAG CGT AAA CGG CTG GTG GCG CAG CAT      576
Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
            180                 185                 190

CTG ACG CAC GTG TCG CAG ACG CAG GAG CAG AAC AAG GAA AAC CAG AAC      624
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
        195                 200                 205

CCC AAT TCT GAC GCG CCG GTC ATC AGG TCA AAA ACC TCC GCC AGG TAC      672
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
    210                 215                 220

ATG GAG CTG GTC GGG TGG CTG GTG GAC CGC GGG ATC ACG TCA GAA AAG      720
Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
225                 230                 235                 240

CAA TGG ATC CAG GAG GAC CAG GCG TCC TAC ATC TCC TTC AAC GCC GCC      768
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
                245                 250                 255

TCC AAC TCG CGG TCA CAA ATC AAG GCC GCG CTG GAC AAT GCC TCC AAA      816
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
            260                 265                 270

ATC ATG AGC CTG ACA AAG ACG GCT CCG GAC TAC CTG GTG GGC CAG AAC      864
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
        275                 280                 285

CCG CCG GAG GAC ATT TCC AGC AAC CGC ATC TAC CGA ATC CTC GAG ATG      912
Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
    290                 295                 300

AAC GGG TAC GAT CCG CAG TAC GCG GCC TCC GTC TTC CTG GGC TGG GCG      960
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305                 310                 315                 320
```

```
CAA AAG AAG TTC GGG AAG AGG AAC ACC ATC TGG CTC TTT GGG CCG GCC      1008
Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
                325                 330                 335

ACG ACG GGT AAA ACC AAC ATC GCG GAA GCC ATC GCC CAC GCC GTG CCC      1056
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
            340                 345                 350

TTC TAC GGC TGC GTG AAC TGG ACC AAT GAG AAC TTT CCG TTC AAC GAT      1104
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
        355                 360                 365

TGC GTC GAC AAG ATG GTG ATC TGG TGG GAG GAG GGC AAG ATG ACG GCC      1152
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
    370                 375                 380

AAG GTC GTA GAG AGC GCC AAG GCC ATC CTG GGC GGA AGC AAG GTG CGC      1200
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385                 390                 395                 400

GTG GAC CAA AAG TGC AAG TCA TCG GCC CAG ATC GAC CCA ACT CCC GTG      1248
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
                405                 410                 415

ATC GTC ACC TCC AAC ACC AAC ATG TGC GCG GTC ATC GAC GGA AAC TCG      1296
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
            420                 425                 430

ACC ACC TTC GAG CAC CAA CAA CCA CTC CAG GAC CGG ATG TTC AAG TTC      1344
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
        435                 440                 445

GAG CTC ACC AAG CGC CTG GAG CAC GAC TTT GGC AAG GTC ACC AAG CAG      1392
Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
    450                 455                 460

GAA GTC AAA GAC TTT TTC CGG TGG GCG TCA GAT CAC GTG ACC GAG GTG      1440
Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465                 470                 475                 480

ACT CAC GAG TTT TAC GTC AGA AAG GGT GGA GCT AGA AAG AGG CCC GCC      1488
Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
                485                 490                 495

CCC AAT GAC GCA GAT ATA AGT GAG CCC AAG CGG GCC TGT CCG TCA GTT      1536
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
            500                 505                 510

GCG CAG CCA TCG ACG TCA GAC GCG GAA GCT CCG GTG GAC TAC GCG GAC      1584
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
        515                 520                 525

AGG TAC CAA AAC AAA TGT TCT CGT CAC GTG GGT ATG AAT CTG ATG CTT      1632
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
    530                 535                 540

TTT CCC TGC CGG CAA TGC GAG AGA ATG AAT CAG AAT GTG GAC ATT TGC      1680
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Val Asp Ile Cys
545                 550                 555                 560

TTC ACG CAC GGG GTC ATG GAC TGT GCC GAG TGC TTC CCC GTG TCA GAA      1728
Phe Thr His Gly Val Met Asp Cys Ala Glu Cys Phe Pro Val Ser Glu
                565                 570                 575
```

```
TCT CAA CCC GTG TCT GTC GTC AGA AAG CGG ACG TAT CAG AAA CTG TGT    1776
Ser Gln Pro Val Ser Val Val Arg Lys Arg Thr Tyr Gln Lys Leu Cys
            580                 585                 590

CCG ATT CAT CAC ATC ATG GGG AGG GCG CCC GAG GTG GCC TGC TCG GCC    1824
Pro Ile His His Ile Met Gly Arg Ala Pro Glu Val Ala Cys Ser Ala
        595                 600                 605

TGC GAA CTG GCC AAT GTG GAC TTG GAT GAC TGT GAC ATG GAA CAA TAA    1872
Cys Glu Leu Ala Asn Val Asp Leu Asp Asp Cys Asp Met Glu Gln  *
    610                 615                 620
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 734 amino acids
(B) TYPE: amino acid

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 capsid protein VP1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Thr Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu
1               5                   10                  15
Gly Val Arg Glu Trp Trp Ala Leu Gln Pro Gly Ala Pro Lys Pro Lys
            20                  25                  30
Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly
        35                  40                  45
Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val
    50                  55                  60
Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln
65                  70                  75                  80
Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp
                85                  90                  95
Ala Glu Phe Gln Gln Arg Leu Gln Gly Asp Thr Ser Phe Gly Gly Asn
            100                 105                 110
Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Leu
        115                 120                 125
Gly Leu Val Glu Gln Ala Gly Glu Thr Ala Pro Gly Lys Lys Arg Pro
    130                 135                 140
Leu Ile Glu Ser Pro Gln Gln Pro Asp Ser Ser Thr Gly Ile Gly Lys
145                 150                 155                 160
Lys Gly Lys Gln Pro Ala Lys Lys Lys Leu Val Phe Glu Asp Glu Thr
                165                 170                 175
Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Thr Ser Gly Ala Met Ser
            180                 185                 190
Asp Asp Ser Glu Met Arg Ala Ala Ala Gly Gly Ala Ala Val Glu Gly
        195                 200                 205
Gly Gln Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
    210                 215                 220
Asp Ser Thr Trp Ser Glu Gly His Val Thr Thr Thr Ser Thr Arg Thr
225                 230                 235                 240
Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Arg Leu Gly Glu
                245                 250                 255
Ser Leu Gln Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
            260                 265                 270
```

Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
275 280 285

Arg Leu Ile Asn Asn Asn Trp Gly Met Arg Pro Lys Ala Met Arg Val
290 295 300

Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305 310 315 320

Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
325 330 335

Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
340 345 350

Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
355 360 365

Cys Gly Leu Val Thr Gly Asn Thr Ser Gln Gln Gln Thr Asp Arg Asn
370 375 380

Ala Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly
385 390 395 400

Asn Asn Phe Glu Ile Thr Tyr Ser Phe Glu Lys Val Pro Phe His Ser
405 410 415

Met Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile
420 425 430

Asp Gln Tyr Leu Trp Gly Leu Gln Ser Thr Thr Thr Gly Thr Thr Leu
435 440 445

Asn Ala Gly Thr Ala Thr Thr Asn Phe Thr Lys Leu Arg Pro Thr Asn
450 455 460

Phe Ser Asn Phe Lys Lys Asn Trp Leu Pro Gly Pro Ser Ile Lys Gln
465 470 475 480

Gln Gly Phe Ser Lys Thr Ala Asn Gln Asn Tyr Lys Ile Pro Ala Thr
485 490 495

Gly Ser Asp Ser Leu Ile Lys Tyr Glu Thr His Ser Thr Leu Asp Gly
500 505 510

Arg Trp Ser Ala Leu Thr Pro Gly Pro Pro Met Ala Thr Ala Gly Pro
515 520 525

Ala Asp Ser Lys Phe Ser Asn Ser Gln Leu Ile Phe Ala Gly Pro Lys
530 535 540

Gln Asn Gly Asn Thr Ala Thr Val Pro Gly Thr Leu Ile Phe Thr Ser
545 550 555 560

Glu Glu Glu Leu Ala Ala Thr Asn Ala Thr Asp Thr Asp Met Trp Gly
565 570 575

Asn Leu Pro Gly Gly Asp Gln Ser Asn Ser Asn Leu Pro Thr Val Asp
580 585 590

Arg Leu Thr Ala Leu Gly Ala Val Pro Gly Met Val Trp Gln Asn Arg
595 600 605

Asp Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp
610 615 620

Gly His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His
625 630 635 640

Pro Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro
645 650 655

Ala Thr Thr Phe Ser Ser Thr Pro Val Asn Ser Phe Ile Thr Gln Tyr
660 665 670

Ser Thr Gly Gln Val Ser Val Gln Ile Asp Trp Glu Ile Gln Lys Glu
675 680 685

Arg Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly
690 695 700

Gln Gln Asn Ser Leu Leu Trp Ala Pro Asp Ala Ala Gly Lys Tyr Thr
705 710 715 720

Glu Pro Arg Ala Ile Gly Thr Arg Tyr Leu Thr His His Leu
725 730

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2208 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 capsid protein VP1 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
ATGACTGACG GTTACCTTCC AGATTGGCTA GAGGACAACC TCTCTGAAGG CGTTCGAGAG     60
TGGTGGGCGC TGCAACCTGG AGCCCCTAAA CCCAAGGCAA ATCAACAACA TCAGGACAAC    120
GCTCGGGGTC TTGTGCTTCC GGGTTACAAA TACCTCGGAC CCGGCAACGG ACTCGACAAG    180
GGGGAACCCG TCAACGCAGC GGACGCGGCA GCCCTCGAGC ACGACAAGGC CTACGACCAG    240
CAGCTCAAGG CCGGTGACAA CCCCTACCTC AAGTACAACC ACGCCGACGC GGAGTTCCAG    300
CAGCGGCTTC AGGGCGACAC ATCGTTTGGG GGCAACCTCG GCAGAGCAGT CTTCCAGGCC    360
AAAAAGAGGG TTCTTGAACC TCTTGGTCTG GTTGAGCAAG CGGGTGAGAC GGCTCCTGGA    420
AAGAAGAGAC CGTTGATTGA ATCCCCCCAG CAGCCCGACT CCTCCACGGG TATCGGCAAA    480
AAAGGCAAGC AGCCGGCTAA AAAGAAGCTC GTTTTCGAAG ACGAAACTGG AGCAGGCGAC    540
GGACCCCCTG AGGGATCAAC TTCCGGAGCC ATGTCTGATG ACAGTGAGAT GCGTGCAGCA    600
GCTGGCGGAG CTGCAGTCGA GGGSGGACAA GGTGCCGATG GAGTGGGTAA TGCCTCGGGT    660
GATTGGCATT GCGATTCCAC CTGGTCTGAG GGCCACGTCA CGACCACCAG CACCAGAACC    720
TGGGTCTTGC CCACCTACAA CAACCACCTN TACAAGCGAC TCGGAGAGAG CCTGCAGTCC    780
AACACCTACA ACGGATTCTC CACCCCCTGG GGATACTTTG ACTTCAACCG CTTCCACTGC    840
CACTTCTCAC CACGTGACTG GCAGCGACTC ATCAACAACA ACTGGGGCAT GCGACCCAAA    900
GCCATGCGGG TCAAAATCTT CAACATCCAG GTCAAGGAGG TCACGACGTC GAACGGCGAG    960
ACAACGGTGG CTAATAACCT TACCAGCACG GTTCAGATCT TTGCGGACTC GTCGTACGAA   1020
CTGCCGTACG TGATGGATGC GGGTCAAGAG GGCAGCCTGC CTCCTTTTCC CAACGACGTC   1080
TTTATGGTGC CCCAGTACGG CTACTGTGGA CTGGTGACCG GCAACACTTC GCAGCAACAG   1140
ACTGACAGAA ATGCCTTCTA CTGCCTGGAG TACTTTCCTT CGCAGATGCT GCGGACTGGC   1200
AACAACTTTG AAATTACGTA CAGTTTTGAG AAGGTGCCTT TCCACTCGAT GTACGCGCAC   1260
AGCCAGAGCC TGGACCGGCT GATGAACCCT CTCATCGACC AGTACCTGTG GGGACTGCAA   1320
TCGACCACCA CCGGAACCAC CCTGAATGCC GGGACTGCCA CCACCAACTT TACCAAGCTG   1380
CGGCCTACCA ACTTTTCCAA CTTTAAAAAG AACTGGCTGC CCGGGCCTTC AATCAAGCAG   1440
CAGGGCTTCT CAAAGACTGC CAATCAAAAC TACAAGATCC CTGCCACCGG GTCAGACAGT   1500
CTCATCAAAT ACGAGACGCA CAGCACTCTG GACGGAAGAT GGAGTGCCCT GACCCCCGGA   1560
CCTCCAATGG CCACGGCTGG ACCTGCGGAC AGCAAGTTCA GCAACAGCCA GCTCATCTTT   1620
GCGGGGCCTA AACAGAACGG CAACACGGCC ACCGTACCCG GGACTCTGAT CTTCACCTCT   1680
GAGGAGGAGC TGGCAGCCAC CAACGCCACC GATACGGACA TGTGGGGCAA CCTACCTGGC   1740
GGTGACCAGA GCAACAGCAA CCTGCCGACC GTGGACAGAC TGACAGCCTT GGGAGCCGTG   1800
CCTGGAATGG TCTGGCAAAA CAGAGACATT TACTACCAGG GTCCCATTTG GGCCAAGATT   1860
CCTCATACCG ATGGACACTT TCACCCCTCA CCGCTGATTG GTGGGTTTGG GCTGAAACAC   1920
CCGCCTCCTC AAATTTTTAT CAAGAACACC CCGGTACCTG CGAATCCTGC AACGACCTTC   1980
AGCTCTACTC CGGTAAACTC CTTCATTACT CAGTACAGCA CTGGCCAGGT GTCGGTGCAG   2040
ATTGACTGGG AGATCCAGAA GGAGCGGTCC AAACGCTGGA ACCCCGAGGT CCAGTTTACC   2100
TCCAACTACG GACAGCAAAA CTCTCTGTTG TGGGCTCCCG ATGCGGCTGG GAAATACACT   2160
GAGCCTAGGG CTATCGGTAC CCGCTACCTC ACCCACCACC TGTAATAA                2208
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 ITR "flip" orientation

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
TTGGCCACTC CCTCTATGCG CGCTCGCTCA CTCACTCGGC CCTGGAGACC AAAGGTCTCC    60
AGACTGCCGG CCTCTGGCCG GCAGGGCCGA GTGAGTGAGC GAGCGCGCAT AGAGGGAGTG   120
GCCAA                                                               125
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 p5 promoter

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
CTCCATCATC TAGGTTTGCC CACTGACGTC AATGTGACGT CCTAGGGTTA GGGAGGTCCC    60
TGTATTAGCA GTCACGTGAG TGTCGTATTT CGCGGAGCGT AGCGGAGCGC ATACCAAGCT   120
GCCACGTCAC AGCCACGTGG TCCGTTTGCG ACAGTTTGCG ACACCATGTG GTCAGGAGGG   180
TATATAACCG CGAGTGAGCC AGCGAGGAGC TCCATTTTGC CCGCGAATTT TGAACGAGCA   240
GCAGC                                                               245
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 Rep protein 40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
1               5                   10                  15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
            20                  25                  30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
        35                  40                  45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
    50                  55                  60
Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
65                  70                  75                  80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
                85                  90                  95
Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
            100                 105                 110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
        115                 120                 125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
    130                 135                 140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
145                 150                 155                 160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
                165                 170                 175
```

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
180 185 190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
195 200 205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
210 215 220
Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
225 230 235 240
Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
245 250 255
Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
260 265 270
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
275 280 285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
290 295 300
Arg Leu Ala Arg Gly Gln Pro Leu Xaa
305 310

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 Rep protein 52

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
1 5 10 15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
20 25 30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
35 40 45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
50 55 60
Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
65 70 75 80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
85 90 95
Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
100 105 110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
115 120 125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
130 135 140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
145 150 155 160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
165 170 175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
180 185 190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
195 200 205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
210 215 220

```
Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
225                 230                 235                 240
Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
                245                 250                 255
Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
            260                 265                 270
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
        275                 280                 285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
    290                 295                 300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
305                 310                 315                 320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Val Asp Ile Cys
                325                 330                 335
Phe Thr His Gly Val Met Asp Cys Ala Glu Cys Phe Pro Val Ser Glu
            340                 345                 350
Ser Gln Pro Val Ser Val Val Arg Lys Arg Thr Tyr Gln Lys Leu Cys
        355                 360                 365
Pro Ile His His Ile Met Gly Arg Ala Pro Glu Val Ala Cys Ser Ala
    370                 375                 380
Cys Glu Leu Ala Asn Val Asp Leu Asp Asp Cys Asp Met Glu Gln
385                 390                 395
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 537 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 Rep protein 68

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Pro Gly Phe Tyr Glu Ile Val Leu Lys Val Pro Ser Asp Leu Asp
1               5                   10                  15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu
            20                  25                  30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
        35                  40                  45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Glu Phe Leu
    50                  55                  60
Val Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65                  70                  75                  80
Gln Phe Glu Lys Gly Asp Ser Tyr Phe His Leu His Ile Leu Val Glu
                85                  90                  95
Thr Val Gly Val Lys Ser Met Val Val Gly Arg Tyr Val Ser Gln Ile
            100                 105                 110
Lys Glu Lys Leu Val Thr Arg Ile Tyr Arg Gly Val Glu Pro Gln Leu
        115                 120                 125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
    130                 135                 140
Asn Lys Val Val Asp Asp Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145                 150                 155                 160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Asp Gln Tyr Ile
                165                 170                 175
Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
            180                 185                 190
```

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195 200 205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210 215 220
Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
225 230 235 240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245 250 255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
260 265 270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
275 280 285
Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
290 295 300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305 310 315 320
Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325 330 335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
340 345 350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355 360 365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370 375 380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385 390 395 400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405 410 415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
420 425 430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435 440 445
Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
450 455 460
Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465 470 475 480
Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
485 490 495
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
500 505 510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
515 520 525
Arg Leu Ala Arg Gly Gln Pro Leu Xaa
530 535

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 623 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 Rep protein 78

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Pro Gly Phe Tyr Glu Ile Val Leu Lys Val Pro Ser Asp Leu Asp
1 5 10 15

```
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu
            20                  25                  30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
        35                  40                  45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Glu Phe Leu
    50                  55                  60
Val Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65                  70                  75                  80
Gln Phe Glu Lys Gly Asp Ser Tyr Phe His Leu His Ile Leu Val Glu
                85                  90                  95
Thr Val Gly Val Lys Ser Met Val Val Gly Arg Tyr Val Ser Gln Ile
            100                 105                 110
Lys Glu Lys Leu Val Thr Arg Ile Tyr Arg Gly Val Glu Pro Gln Leu
        115                 120                 125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
    130                 135                 140
Asn Lys Val Val Asp Asp Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145                 150                 155                 160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Asp Gln Tyr Ile
                165                 170                 175
Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
            180                 185                 190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
        195                 200                 205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
    210                 215                 220
Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
225                 230                 235                 240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
                245                 250                 255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Ser Lys
            260                 265                 270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Asn
        275                 280                 285
Pro Pro Glu Asp Ile Ser Ser Asn Arg Ile Tyr Arg Ile Leu Glu Met
    290                 295                 300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305                 310                 315                 320
Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
                325                 330                 335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
            340                 345                 350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
        355                 360                 365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
    370                 375                 380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385                 390                 395                 400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
                405                 410                 415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
            420                 425                 430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
        435                 440                 445
Glu Leu Thr Lys Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
    450                 455                 460
Glu Val Lys Asp Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465                 470                 475                 480
Thr His Glu Phe Tyr Val Arg Lys Gly Gly Ala Arg Lys Arg Pro Ala
                485                 490                 495
Pro Asn Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
            500                 505                 510
```

```
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Pro Val Asp Tyr Ala Asp
        515                 520                 525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
    530                 535                 540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Val Asp Ile Cys
545                 550                 555                 560
Phe Thr His Gly Val Met Asp Cys Ala Glu Cys Phe Pro Val Ser Glu
                565                 570                 575
Ser Gln Pro Val Ser Val Val Arg Lys Arg Thr Tyr Gln Lys Leu Cys
            580                 585                 590
Pro Ile His His Ile Met Gly Arg Ala Pro Glu Val Ala Cys Ser Ala
        595                 600                 605
Cys Glu Leu Ala Asn Val Asp Leu Asp Asp Cys Asp Met Glu Gln
    610                 615                 620
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 939 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 Rep 40 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
ATGGAGCTGG TCGGGTGGCT GGTGGACCGC GGGATCACGT CAGAAAAGCA ATGGATCCAG    60
GAGGACCAGG CGTCCTACAT CTCCTTCAAC GCCGCCTCCA ACTCGCGGTC ACAAATCAAG   120
GCCGCGCTGG ACAATGCCTC CAAAATCATG AGCCTGACAA AGACGGCTCC GGACTACCTG   180
GTGGGCCAGA ACCCGCCGGA GGACATTTCC AGCAACCGCA TCTACCGAAT CCTCGAGATG   240
AACGGGTACG ATCCGCAGTA CGCGGCCTCC GTCTTCCTGG GCTGGGCGCA AAAGAAGTTC   300
GGGAAGAGGA ACACCATCTG GCTCTTTGGG CCGGCCACGA CGGGTAAAAC CAACATCGCG   360
GAAGCCATCG CCCACGCCGT GCCCTTCTAC GGCTGCGTGA ACTGGACCAA TGAGAACTTT   420
CCGTTCAACG ATTGCGTCGA CAAGATGGTG ATCTGGTGGG AGGAGGGCAA GATGACGGCC   480
AAGGTCGTAG AGAGCGCCAA GGCCATCCTG GGCGGAAGCA AGGTGCGCGT GGACCAAAAG   540
TGCAAGTCAT CGGCCCAGAT CGACCCAACT CCCGTGATCG TCACCTCCAA CACCAACATG   600
TGCGCGGTCA TCGACGGAAA CTCGACCACC TTCGAGCACC AACAACCACT CCAGGACCGG   660
ATGTTCAAGT TCGAGCTCAC CAAGCGCCTG GAGCACGACT TTGGCAAGGT CACCAAGCAG   720
GAAGTCAAAG ACTTTTTCCG GTGGGCGTCA GATCACGTGA CCGAGGTGAC TCACGAGTTT   780
TACGTCAGAA AGGGTGGAGC TAGAAAGAGG CCCGCCCCCA ATGACGCAGA TATAAGTGAG   840
CCCAAGCGGG CCTGTCCGTC AGTTGCGCAG CCATCGACGT CAGACGCGGA AGCTCCGGTG   900
GACTACGCGG ACAGATTGGC TAGAGGACAA CCTCTCTGA                          939
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1197 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 Rep 52 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
ATGGAGCTGG TCGGGTGGCT GGTGGACCGC GGGATCACGT CAGAAAAGCA ATGGATCCAG    60
GAGGACCAGG CGTCCTACAT CTCCTTCAAC GCCGCCTCCA ACTCGCGGTC ACAAATCAAG   120
GCCGCGCTGG ACAATGCCTC CAAAATCATG AGCCTGACAA AGACGGCTCC GGACTACCTG   180
GTGGGCCAGA ACCCGCCGGA GGACATTTCC AGCAACCGCA TCTACCGAAT CCTCGAGATG   240
```

AACGGGTACG ATCCGCAGTA CGCGGCCTCC GTCTTCCTGG GCTGGGCGCA AAAGAAGTTC 300
GGGAAGAGGA ACACCATCTG GCTCTTTGGG CCGGCCACGA CGGGTAAAAC CAACATCGCG 360
GAAGCCATCG CCCACGCCGT GCCCTTCTAC GGCTGCGTGA ACTGGACCAA TGAGAACTTT 420
CCGTTCAACG ATTGCGTCGA CAAGATGGTG ATCTGGTGGG AGGAGGGCAA GATGACGGCC 480
AAGGTCGTAG AGAGCGCCAA GGCCATCCTG GGCGGAAGCA AGGTGCGCGT GGACCAAAAG 540
TGCAAGTCAT CGGCCCAGAT CGACCCAACT CCCGTGATCG TCACCTCCAA CACCAACATG 600
TGCGCGGTCA TCGACGGAAA CTCGACCACC TTCGAGCACC AACAACCACT CCAGGACCGG 660
ATGTTCAAGT TCGAGCTCAC CAAGCGCCTG GAGCACGACT TTGGCAAGGT CACCAAGCAG 720
GAAGTCAAAG ACTTTTTCCG GTGGGCGTCA GATCACGTGA CCGAGGTGAC TCACGAGTTT 780
TACGTCAGAA AGGGTGGAGC TAGAAAGAGG CCCGCCCCCA ATGACGCAGA TATAAGTGAG 840
CCCAAGCGGG CCTGTCCGTC AGTTGCGCAG CCATCGACGT CAGACGCGGA AGCTCCGGTG 900
GACTACGCGG ACAGGTACCA AAACAAATGT TCTCGTCACG TGGGTATGAA TCTGATGCTT 960
TTTCCCTGCC GGCAATGCGA GAGAATGAAT CAGAATGTGG ACATTTGCTT CACGCACGGG 1020
GTCATGGACT GTGCCGAGTG CTTCCCCGTG TCAGAATCTC AACCCGTGTC TGTCGTCAGA 1080
AAGCGGACGT ATCAGAAACT GTGTCCGATT CATCACATCA TGGGGAGGGC GCCCGAGGTG 1140
GCCTGCTCGG CCTGCGAACT GGCCAATGTG GACTTGGATG ACTGTGACAT GGAACAA 1197

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1611 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 Rep 68 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGCCGGGGT TCTACGAGAT CGTGCTGAAG GTGCCCAGCG ACCTGGACGA GCACCTGCCC 60
GGCATTTCTG ACTCTTTTGT GAGCTGGGTG GCCGAGAAGG AATGGGAGCT GCCGCCGGAT 120
TCTGACATGG ACTTGAATCT GATTGAGCAG GCACCCCTGA CCGTGGCCGA AAAGCTGCAA 180
CGCGAGTTCC TGGTCGAGTG GCGCCGCGTG AGTAAGGCCC CGGAGGCCCT CTTCTTTGTC 240
CAGTTCGAGA AGGGGGACAG CTACTTCCAC CTGCACATCC TGGTGGAGAC CGTGGGCGTC 300
AAATCCATGG TGGTGGGCCG CTACGTGAGC CAGATTAAAG AGAAGCTGGT GACCCGCATC 360
TACCGCGGGG TCGAGCCGCA GCTTCCGAAC TGGTTCGCGG TGACCAAGAC GCGTAATGGC 420
GCCGGAGGCG GGAACAAGGT GGTGGACGAC TGCTACATCC CCAACTACCT GCTCCCCAAG 480
ACCCAGCCCG AGCTCCAGTG GGCGTGGACT AACATGGACC AGTATATAAG CGCCTGTTTG 540
AATCTCGCGG AGCGTAAACG GCTGGTGGCG CAGCATCTGA CGCACGTGTC GCAGACGCAG 600
GAGCAGAACA AGGAAAACCA GAACCCCAAT TCTGACGCGC CGGTCATCAG GTCAAAAACC 660
TCCGCCAGGT ACATGGAGCT GGTCGGGTGG CTGGTGGACC GCGGGATCAC GTCAGAAAAG 720
CAATGGATCC AGGAGGACCA GGCGTCCTAC ATCTCCTTCA ACGCCGCCTC CAACTCGCGG 780
TCACAAATCA AGGCCGCGCT GGACAATGCC TCCAAAATCA TGAGCCTGAC AAAGACGGCT 840
CCGGACTACC TGGTGGGCCA GAACCCGCCG GAGGACATTT CCAGCAACCG CATCTACCGA 900
ATCCTCGAGA TGAACGGGTA CGATCCGCAG TACGCGGCCT CCGTCTTCCT GGGCTGGGCG 960
CAAAAGAAGT TCGGGAAGAG GAACACCATC TGGCTCTTTG GGCCGGCCAC GACGGGTAAA 1020
ACCAACATCG CGGAAGCCAT CGCCCACGCC GTGCCCTTCT ACGGCTGCGT GAACTGGACC 1080
AATGAGAACT TTCCGTTCAA CGATTGCGTC GACAAGATGG TGATCTGGTG GGAGGAGGGC 1140
AAGATGACGG CCAAGGTCGT AGAGAGCGCC AAGGCCATCC TGGGCGGAAG CAAGGTGCGC 1200
GTGGACCAAA AGTGCAAGTC ATCGGCCCAG ATCGACCCAA CTCCCGTGAT CGTCACCTCC 1260
AACACCAACA TGTGCGCGGT CATCGACGGA AACTCGACCA CCTTCGAGCA CCAACAACCA 1320
CTCCAGGACC GGATGTTCAA GTTCGAGCTC ACCAAGCGCC TGGAGCACGA CTTTGGCAAG 1380
GTCACCAAGC AGGAAGTCAA AGACTTTTTC CGGTGGGCGT CAGATCACGT GACCGAGGTG 1440
ACTCACGAGT TTTACGTCAG AAAGGGTGGA GCTAGAAAGA GGCCCGCCCC CAATGACGCA 1500
GATATAAGTG AGCCCAAGCG GGCCTGTCCG TCAGTTGCGC AGCCATCGAC GTCAGACGCG 1560
GAAGCTCCGG TGGACTACGC GGACAGATTG GCTAGAGGAC AACCTCTCTG A 1611

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1872 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 Rep 78 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
ATGCCGGGGT TCTACGAGAT CGTGCTGAAG GTGCCCAGCG ACCTGGACGA GCACCTGCCC     60
GGCATTTCTG ACTCTTTTGT GAGCTGGGTG GCCGAGAAGG AATGGGAGCT GCCGCCGGAT    120
TCTGACATGG ACTTGAATCT GATTGAGCAG GCACCCCTGA CCGTGGCCGA AAAGCTGCAA    180
CGCGAGTTCC TGGTCGAGTG GCGCCGCGTG AGTAAGGCCC CGGAGGCCCT CTTCTTTGTC    240
CAGTTCGAGA AGGGGGACAG CTACTTCCAC CTGCACATCC TGGTGGAGAC CGTGGGCGTC    300
AAATCCATGG TGGTGGGCCG CTACGTGAGC CAGATTAAAG AGAAGCTGGT GACCCGCATC    360
TACCGCGGGG TCGAGCCGCA GCTTCCGAAC TGGTTCGCGG TGACCAAGAC GCGTAATGGC    420
GCCGGAGGCG GGAACAAGGT GGTGGACGAC TGCTACATCC CCAACTACCT GCTCCCCAAG    480
ACCCAGCCCG AGCTCCAGTG GGCGTGGACT AACATGGACC AGTATATAAG CGCCTGTTTG    540
AATCTCGCGG AGCGTAAACG GCTGGTGGCG CAGCATCTGA CGCACGTGTC GCAGACGCAG    600
GAGCAGAACA AGGAAAACCA GAACCCCAAT TCTGACGCGC CGGTCATCAG GTCAAAAACC    660
TCCGCCAGGT ACATGGAGCT GGTCGGGTGG CTGGTGGACC GCGGGATCAC GTCAGAAAAG    720
CAATGGATCC AGGAGGACCA GGCGTCCTAC ATCTCCTTCA ACGCCGCCTC CAACTCGCGG    780
TCACAAATCA AGGCCGCGCT GGACAATGCC TCCAAAATCA TGAGCCTGAC AAAGACGGCT    840
CCGGACTACC TGGTGGGCCA GAACCCGCCG GAGGACATTT CCAGCAACCG CATCTACCGA    900
ATCCTCGAGA TGAACGGGTA CGATCCGCAG TACGCGGCCT CCGTCTTCCT GGGCTGGGCG    960
CAAAAGAAGT TCGGGAAGAG GAACACCATC TGGCTCTTTG GGCCGGCCAC GACGGGTAAA   1020
ACCAACATCG CGGAAGCCAT CGCCCACGCC GTGCCCTTCT ACGGCTGCGT GAACTGGACC   1080
AATGAGAACT TTCCGTTCAA CGATTGCGTC GACAAGATGG TGATCTGGTG GGAGGAGGGC   1140
AAGATGACGG CCAAGGTCGT AGAGAGCGCC AAGGCCATCC TGGGCGGAAG CAAGGTGCGC   1200
GTGGACCAAA AGTGCAAGTC ATCGGCCCAG ATCGACCCAA CTCCCGTGAT CGTCACCTCC   1260
AACACCAACA TGTGCGCGGT CATCGACGGA AACTCGACCA CCTTCGAGCA CCAACAACCA   1320
CTCCAGGACC GGATGTTCAA GTTCGAGCTC ACCAAGCGCC TGGAGCACGA CTTTGGCAAG   1380
GTCACCAAGC AGGAAGTCAA AGACTTTTTC CGGTGGGCGT CAGATCACGT GACCGAGGTG   1440
ACTCACGAGT TTTACGTCAG AAAGGGTGGA GCTAGAAAGA GGCCCGCCCC CAATGACGCA   1500
GATATAAGTG AGCCCAAGCG GGCCTGTCCG TCAGTTGCGC AGCCATCGAC GTCAGACGCG   1560
GAAGCTCCGG TGGACTACGC GGACAGGTAC CAAAACAAAT GTTCTCGTCA CGTGGGTATG   1620
AATCTGATGC TTTTTCCCTG CCGGCAATGC GAGAGAATGA ATCAGAATGT GGACATTTGC   1680
TTCACGCACG GGGTCATGGA CTGTGCCGAG TGCTTCCCCG TGTCAGAATC TCAACCCGTG   1740
TCTGTCGTCA GAAAGCGGAC GTATCAGAAA CTGTGTCCGA TTCATCACAT CATGGGGAGG   1800
GCGCCCGAGG TGGCCTGCTC GGCCTGCGAA CTGGCCAATG TGGACTTGGA TGACTGTGAC   1860
ATGGAACAAT AA                                                       1872
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 598 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 capsid protein VP2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Thr Ala Pro Gly Lys Lys Arg Pro Leu Ile Glu Ser Pro Gln Gln Pro
 1               5                  10                  15
Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Lys Gln Pro Ala Lys Lys
            20                  25                  30
```

Lys Leu Val Phe Glu Asp Glu Thr Gly Ala Gly Asp Gly Pro Pro Glu
35 40 45
Gly Ser Thr Ser Gly Ala Met Ser Asp Asp Ser Glu Met Arg Ala Ala
50 55 60
Ala Gly Gly Ala Ala Val Glu Gly Gly Gln Gly Ala Asp Gly Val Gly
65 70 75 80
Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr Trp Ser Glu Gly His
85 90 95
Val Thr Thr Thr Ser Thr Arg Thr Trp Val Leu Pro Thr Tyr Asn Asn
100 105 110
His Leu Tyr Lys Arg Leu Gly Glu Ser Leu Gln Ser Asn Thr Tyr Asn
115 120 125
Gly Phe Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys
130 135 140
His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly
145 150 155 160
Met Arg Pro Lys Ala Met Arg Val Lys Ile Phe Asn Ile Gln Val Lys
165 170 175
Glu Val Thr Thr Ser Asn Gly Glu Thr Thr Val Ala Asn Asn Leu Thr
180 185 190
Ser Thr Val Gln Ile Phe Ala Asp Ser Ser Tyr Glu Leu Pro Tyr Val
195 200 205
Met Asp Ala Gly Gln Glu Gly Ser Leu Pro Pro Phe Pro Asn Asp Val
210 215 220
Phe Met Val Pro Gln Tyr Gly Tyr Cys Gly Leu Val Thr Gly Asn Thr
225 230 235 240
Ser Gln Gln Gln Thr Asp Arg Asn Ala Phe Tyr Cys Leu Glu Tyr Phe
245 250 255
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Ile Thr Tyr Ser
260 265 270
Phe Glu Lys Val Pro Phe His Ser Met Tyr Ala His Ser Gln Ser Leu
275 280 285
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Trp Gly Leu Gln
290 295 300
Ser Thr Thr Thr Gly Thr Thr Leu Asn Ala Gly Thr Ala Thr Thr Asn
305 310 315 320
Phe Thr Lys Leu Arg Pro Thr Asn Phe Ser Asn Phe Lys Lys Asn Trp
325 330 335
Leu Pro Gly Pro Ser Ile Lys Gln Gln Gly Phe Ser Lys Thr Ala Asn
340 345 350
Gln Asn Tyr Lys Ile Pro Ala Thr Gly Ser Asp Ser Leu Ile Lys Tyr
355 360 365
Glu Thr His Ser Thr Leu Asp Gly Arg Trp Ser Ala Leu Thr Pro Gly
370 375 380
Pro Pro Met Ala Thr Ala Gly Pro Ala Asp Ser Lys Phe Ser Asn Ser
385 390 395 400
Gln Leu Ile Phe Ala Gly Pro Lys Gln Asn Gly Asn Thr Ala Thr Val
405 410 415
Pro Gly Thr Leu Ile Phe Thr Ser Glu Glu Glu Leu Ala Ala Thr Asn
420 425 430
Ala Thr Asp Thr Asp Met Trp Gly Asn Leu Pro Gly Gly Asp Gln Ser
435 440 445
Asn Ser Asn Leu Pro Thr Val Asp Arg Leu Thr Ala Leu Gly Ala Val
450 455 460
Pro Gly Met Val Trp Gln Asn Arg Asp Ile Tyr Tyr Gln Gly Pro Ile
465 470 475 480
Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro Leu
485 490 495
Ile Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Phe Ile Lys
500 505 510
Asn Thr Pro Val Pro Ala Asn Pro Ala Thr Thr Phe Ser Ser Thr Pro
515 520 525

```
Val Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Gln
    530                 535                 540
Ile Asp Trp Glu Ile Gln Lys Glu Arg Ser Lys Arg Trp Asn Pro Glu
545                 550                 555                 560
Val Gln Phe Thr Ser Asn Tyr Gly Gln Gln Asn Ser Leu Leu Trp Ala
                565                 570                 575
Pro Asp Ala Ala Gly Lys Tyr Thr Glu Pro Arg Ala Ile Gly Thr Arg
            580                 585                 590
Tyr Leu Thr His His Leu
        595
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1800 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 capsid protein VP2 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
ACGGCTCCTG GAAAGAAGAG ACCGTTGATT GAATCCCCCC AGCAGCCCGA CTCCTCCACG     60
GGTATCGGCA AAAAAGGCAA GCAGCCGGCT AAAAAGAAGC TCGTTTTCGA AGACGAAACT    120
GGAGCAGGCG ACGGACCCCC TGAGGGATCA ACTTCCGGAG CCATGTCTGA TGACAGTGAG    180
ATGCGTGCAG CAGCTGGCGG AGCTGCAGTC GAGGGSGGAC AAGGTGCCGA TGGAGTGGGT    240
AATGCCTCGG GTGATTGGCA TTGCGATTCC ACCTGGTCTG AGGGCCACGT CACGACCACC    300
AGCACCAGAA CCTGGGTCTT GCCCACCTAC AACAACCACC TNTACAAGCG ACTCGGAGAG    360
AGCCTGCAGT CCAACACCTA CAACGGATTC TCCACCCCCT GGGGATACTT TGACTTCAAC    420
CGCTTCCACT GCCACTTCTC ACCACGTGAC TGGCAGCGAC TCATCAACAA CAACTGGGGC    480
ATGCGACCCA AAGCCATGCG GGTCAAAATC TTCAACATCC AGGTCAAGGA GGTCACGACG    540
TCGAACGGCG AGACAACGGT GGCTAATAAC CTTACCAGCA CGGTTCAGAT CTTTGCGGAC    600
TCGTCGTACG AACTGCCGTA CGTGATGGAT GCGGGTCAAG AGGGCAGCCT GCCTCCTTTT    660
CCCAACGACG TCTTTATGGT GCCCCAGTAC GGCTACTGTG GACTGGTGAC CGGCAACACT    720
TCGCAGCAAC AGACTGACAG AAATGCCTTC TACTGCCTGG AGTACTTTCC TTCGCAGATG    780
CTGCGGACTG GCAACAACTT TGAAATTACG TACAGTTTTG AGAAGGTGCC TTTCCACTCG    840
ATGTACGCGC ACAGCCAGAG CCTGGACCGG CTGATGAACC CTCTCATCGA CCAGTACCTG    900
TGGGGACTGC AATCGACCAC CACCGGAACC ACCCTGAATG CCGGGACTGC CACCACCAAC    960
TTTACCAAGC TGCGGCCTAC CAACTTTTCC AACTTTAAAA AGAACTGGCT GCCCGGGCCT   1020
TCAATCAAGC AGCAGGGCTT CTCAAAGACT GCCAATCAAA ACTACAAGAT CCCTGCCACC   1080
GGGTCAGACA GTCTCATCAA ATACGAGACG CACAGCACTC TGGACGGAAG ATGGAGTGCC   1140
CTGACCCCCG GACCTCCAAT GGCCACGGCT GGACCTGCGG ACAGCAAGTT CAGCAACAGC   1200
CAGCTCATCT TTGCGGGGCC TAAACAGAAC GGCAACACGG CCACCGTACC CGGGACTCTG   1260
ATCTTCACCT CTGAGGAGGA GCTGGCAGCC ACCAACGCCA CCGATACGGA CATGTGGGGC   1320
AACCTACCTG GCGGTGACCA GAGCAACAGC AACCTGCCGA CCGTGGACAG ACTGACAGCC   1380
TTGGGAGCCG TGCCTGGAAT GGTCTGGCAA AACAGAGACA TTTACTACCA GGGTCCCATT   1440
TGGGCCAAGA TTCCTCATAC CGATGGACAC TTTCACCCCT CACCGCTGAT TGGTGGGTTT   1500
GGGCTGAAAC ACCCGCCTCC TCAAATTTTT ATCAAGAACA CCCCGGTACC TGCGAATCCT   1560
GCAACGACCT TCAGCTCTAC TCCGGTAAAC TCCTTCATTA CTCAGTACAG CACTGGCCAG   1620
GTGTCGGTGC AGATTGACTG GGAGATCCAG AAGGAGCGGT CCAAACGCTG GAACCCCGAG   1680
GTCCAGTTTA CCTCCAACTA CGGACAGCAA AACTCTCTGT TGTGGGCTCC CGATGCGGCT   1740
GGGAAATACA CTGAGCCTAG GGCTATCGGT ACCCGCTACC TCACCCACCA CCTGTAATAA   1800
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 544 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(ix) OTHER INFO: AAV4 capsid protein VP3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met Ser Asp Asp Ser Glu Met Arg Ala Ala Ala Gly Gly Ala Ala Val
1               5                   10                  15
Glu Gly Gly Gln Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp
            20                  25                  30
His Cys Asp Ser Thr Trp Ser Glu Gly His Val Thr Thr Thr Ser Thr
        35                  40                  45
Arg Thr Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Arg Leu
    50                  55                  60
Gly Glu Ser Leu Gln Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp
65                  70                  75                  80
Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp
                85                  90                  95
Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Met Arg Pro Lys Ala Met
            100                 105                 110
Arg Val Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn
        115                 120                 125
Gly Glu Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe
    130                 135                 140
Ala Asp Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu
145                 150                 155                 160
Gly Ser Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr
                165                 170                 175
Gly Tyr Cys Gly Leu Val Thr Gly Asn Thr Ser Gln Gln Gln Thr Asp
            180                 185                 190
Arg Asn Ala Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
        195                 200                 205
Thr Gly Asn Asn Phe Glu Ile Thr Tyr Ser Phe Glu Lys Val Pro Phe
    210                 215                 220
His Ser Met Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
225                 230                 235                 240
Leu Ile Asp Gln Tyr Leu Trp Gly Leu Gln Ser Thr Thr Thr Gly Thr
                245                 250                 255
Thr Leu Asn Ala Gly Thr Ala Thr Thr Asn Phe Thr Lys Leu Arg Pro
            260                 265                 270
Thr Asn Phe Ser Asn Phe Lys Lys Asn Trp Leu Pro Gly Pro Ser Ile
        275                 280                 285
Lys Gln Gln Gly Phe Ser Lys Thr Ala Asn Gln Asn Tyr Lys Ile Pro
    290                 295                 300
Ala Thr Gly Ser Asp Ser Leu Ile Lys Tyr Glu Thr His Ser Thr Leu
305                 310                 315                 320
Asp Gly Arg Trp Ser Ala Leu Thr Pro Gly Pro Pro Met Ala Thr Ala
                325                 330                 335
Gly Pro Ala Asp Ser Lys Phe Ser Asn Ser Gln Leu Ile Phe Ala Gly
            340                 345                 350
Pro Lys Gln Asn Gly Asn Thr Ala Thr Val Pro Gly Thr Leu Ile Phe
        355                 360                 365
Thr Ser Glu Glu Glu Leu Ala Ala Thr Asn Ala Thr Asp Thr Asp Met
    370                 375                 380
Trp Gly Asn Leu Pro Gly Gly Asp Gln Ser Asn Ser Asn Leu Pro Thr
385                 390                 395                 400
Val Asp Arg Leu Thr Ala Leu Gly Ala Val Pro Gly Met Val Trp Gln
                405                 410                 415
```

Asn Arg Asp Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His
420 425 430

Thr Asp Gly His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu
435 440 445

Lys His Pro Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala
450 455 460

Asn Pro Ala Thr Thr Phe Ser Ser Thr Pro Val Asn Ser Phe Ile Thr
465 470 475 480

Gln Tyr Ser Thr Gly Gln Val Ser Val Gln Ile Asp Trp Glu Ile Gln
485 490 495

Lys Glu Arg Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn
500 505 510

Tyr Gly Gln Gln Asn Ser Leu Leu Trp Ala Pro Asp Ala Ala Gly Lys
515 520 525

Tyr Thr Glu Pro Arg Ala Ile Gly Thr Arg Tyr Leu Thr His His Leu
530 535 540

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1617 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 capsid protein VP3 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGCGTGCAG CAGCTGGCGG AGCTGCAGTC GAGGGSGGAC AAGGTGCCGA TGGAGTGGGT 60
AATGCCTCGG GTGATTGGCA TTGCGATTCC ACCTGGTCTG AGGGCCACGT CACGACCACC 120
AGCACCAGAA CCTGGGTCTT GCCCACCTAC AACAACCACC TNTACAAGCG ACTCGGAGAG 180
AGCCTGCAGT CCAACACCTA CAACGGATTC TCCACCCCCT GGGGATACTT TGACTTCAAC 240
CGCTTCCACT GCCACTTCTC ACCACGTGAC TGGCAGCGAC TCATCAACAA CAACTGGGGC 300
ATGCGACCCA AAGCCATGCG GGTCAAAATC TTCAACATCC AGGTCAAGGA GGTCACGACG 360
TCGAACGGCG AGACAACGGT GGCTAATAAC CTTACCAGCA CGGTTCAGAT CTTTGCGGAC 420
TCGTCGTACG AACTGCCGTA CGTGATGGAT GCGGGTCAAG AGGGCAGCCT GCCTCCTTTT 480
CCCAACGACG TCTTTATGGT GCCCCAGTAC GGCTACTGTG GACTGGTGAC CGGCAACACT 540
TCGCAGCAAC AGACTGACAG AAATGCCTTC TACTGCCTGG AGTACTTTCC TTCGCAGATG 600
CTGCGGACTG GCAACAACTT TGAAATTACG TACAGTTTTG AGAAGGTGCC TTTCCACTCG 660
ATGTACGCGC ACAGCCAGAG CCTGGACCGG CTGATGAACC CTCTCATCGA CCAGTACCTG 720
TGGGGACTGC AATCGACCAC CACCGGAACC ACCCTGAATG CCGGGACTGC CACCACCAAC 780
TTTACCAAGC TGCGGCCTAC CAACTTTTCC AACTTTAAAA AGAACTGGCT GCCCGGGCCT 840
TCAATCAAGC AGCAGGGCTT CTCAAAGACT GCCAATCAAA ACTACAAGAT CCCTGCCACC 900
GGGTCAGACA GTCTCATCAA ATACGAGACG CACAGCACTC TGGACGGAAG ATGGAGTGCC 960
CTGACCCCCG GACCTCCAAT GGCCACGGCT GGACCTGCGG ACAGCAAGTT CAGCAACAGC 1020
CAGCTCATCT TTGCGGGGCC TAAACAGAAC GGCAACACGG CCACCGTACC CGGGACTCTG 1080
ATCTTCACCT CTGAGGAGGA GCTGGCAGCC ACCAACGCCA CCGATACGGA CATGTGGGGC 1140
AACCTACCTG GCGGTGACCA GAGCAACAGC AACCTGCCGA CCGTGGACAG ACTGACAGCC 1200
TTGGGAGCCG TGCCTGGAAT GGTCTGGCAA AACAGAGACA TTTACTACCA GGGTCCCATT 1260
TGGGCCAAGA TTCCTCATAC CGATGGACAC TTTCACCCCT CACCGCTGAT TGGTGGGTTT 1320
GGGCTGAAAC ACCCGCCTCC TCAAATTTTT ATCAAGAACA CCCCGGTACC TGCGAATCCT 1380
GCAACGACCT TCAGCTCTAC TCCGGTAAAC TCCTTCATTA CTCAGTACAG CACTGGCCAG 1440
GTGTCGGTGC AGATTGACTG GGAGATCCAG AAGGAGCGGT CCAAACGCTG GAACCCCGAG 1500
GTCCAGTTTA CCTCCAACTA CGGACAGCAA AACTCTCTGT TGTGGCTCC CGATGCGGCT 1560
GGGAAATACA CTGAGCCTAG GGCTATCGGT ACCCGCTACC TCACCCACCA CCTGTAA 1617

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 129 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) OTHER INFO: AAV4 ITR "flop" orientation

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTGGCCACTC CCTCTATGCG CGCTCGCTCA CTCACTCGGC CCTGCGGCCA GAGGCCGGCA 60
GTCTGGAGAC CTTTGGTGTC CAGGGCAGGG CCGAGTGAGT GAGCGAGCGC GCATAGAGGG 120
AGTGGCCAA 129

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TCTAGTCTAG ACTTGGCCAC TCCCTCTCTG CGCGC 35

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGGCCTTAAG AGCAGTCGTC CACCACCTTG TTCC 34

What is claimed is:

1. A nucleic acid vector comprising a pair of adeno-associated virus 4 (AAV4) inverted terminal repeats and a promoter between the inverted terminal repeats.

2. The vector of claim 1, wherein the AAV4 inverted terminal repeats comprise the nucleotide sequence set forth in SEQ ID NO: 6.

3. The vector of claim 1, wherein the AAV4 inverted terminal repeats comprise the nucleotide sequence set forth in SEQ ID NO: 20.

4. The vector of claim 1, wherein the promoter is an AAV promoter p5.

5. The vector of claim 1, wherein the p5 promoter is AAV4 p5 promoter.

6. The vector of claim 1, further comprising an exogenous nucleic acid functionally linked to the promoter.

7. The vector of claim 1 encapsidated in an adeno-associated virus particle.

8. The particle of claim 7, wherein the particle is an AAV4 particle.

9. The particle of claim 7, wherein the particle is an AAV1 particle, an AAV2 particle, an AAV3 particle or an AAV5 particle.

10. An AAV4 particle containing a vector comprising a pair of AAV2 inverted terminal repeats.

11. The particle of claim 10, wherein the vector further comprises an exogenous nucleic acid inserted between the inverted terminal repeats.

12. An isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1.

13. An isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:1.

14. An isolated nucleic acid that selectively hybridizes with the nucleic acid of claim 13.

15. An isolated nucleic acid encoding an adeno-associated virus 4 Rep protein.

16. The nucleic acid of claim 15, wherein the adeno-associated virus 4 Rep protein has the amino acid sequence set forth in SEQ ID NO:2.

17. The nucleic acid of claim 15, wherein the adeno-associated virus 4 Rep protein has the amino acid sequence set forth in SEQ ID NO:8.

18. The nucleic acid of claim 15, wherein the adeno-associated virus 4 Rep protein has the amino acid sequence set forth in SEQ ID NO:9.

19. The nucleic acid of claim 15, wherein the adeno-associated virus 4 Rep protein has the amino acid sequence set forth in SEQ ID NO:10.

20. The nucleic acid of claim 15, wherein the adeno-associated virus 4 Rep protein has the amino acid sequence set forth in SEQ ID NO:11.

21. The nucleic acid of claim 15, wherein the nucleic acid comprises the nucleotide sequence set forth in SEQ ID NO:3.

22. The nucleic acid of claim 15, wherein the nucleic acid consists essentially of the nucleotide sequence set forth in SEQ ID NO:3.

23. An isolated nucleic acid that selectively hybridizes with the nucleic acid of claim 22.

24. The nucleic acid of claim 15, wherein the nucleic acid comprises the nucleotide sequence set forth in SEQ ID NO:12.

25. The nucleic acid of claim 15, wherein the nucleic acid comprises the nucleotide sequence set forth in SEQ ID NO:13.

26. The nucleic acid of claim 15, wherein the nucleic acid comprises the nucleotide sequence set forth in SEQ ID NO:14.

27. The nucleic acid of claim 15, wherein the nucleic acid comprises the nucleotide sequence set forth in SEQ ID NO:15.

28. An isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:2, or a unique fragment thereof.

29. An isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:8, or a unique fragment thereof.

30. An isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:9, or a unique fragment thereof.

31. An isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:10, or a unique fragment thereof.

32. An isolated AAV4 Rep protein having the amino acid sequence set forth in SEQ ID NO:11, or a unique fragment thereof.

33. An isolated antibody that specifically binds the protein of claim 28.

34. An isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:4.

35. An isolated antibody that specifically binds the protein of claim 34.

36. An isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:16.

37. An isolated antibody that specifically binds the protein of claim 36.

38. An isolated AAV4 capsid protein having the amino acid sequence set forth in SEQ ID NO:18.

39. An isolated antibody that specifically binds the protein of claim 38.

40. An isolated nucleic acid encoding adeno-associated virus 4 capsid protein.

41. An isolated nucleic acid encoding the protein of claim 34.

42. The nucleic acid of claim 41, wherein the nucleic acid comprises the nucleic acid sequence set forth in SEQ ID NO:5.

43. The nucleic acid of claim 41, wherein the nucleic acid consists essentially of the nucleic acid sequence set forth in SEQ ID NO:5.

44. An isolated nucleic acid that selectively hybridizes with the nucleic acid of claim 39.

45. An AAV4 particle comprising a capsid protein consisting essentially of the amino acid sequence set forth in SEQ ID NO:4.

46. An isolated nucleic acid comprising an AAV4 p5 promoter.

47. A method of screening a cell for infectivity by AAV4 comprising contacting the cell with AAV4 and detecting the presence of AAV4 in the cells.

48. The method of claim 47, wherein the presence of AAV4 is detected in the cells by nucleic acid hybridization.

49. A method of determining the suitability of an AAV4 vector for administration to a subject comprising administering to an antibody-containing sample from the subject an antigenic fragment of the protein of claim 37 and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the AAV4 vector to be unsuitable for use in the subject.

50. A method of determining the suitability of an AAV4 vector for administration to a subject comprising administering to an antibody-containing sample from the subject an antigenic fragment of the protein of claim 15 and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the AAV4 vector to be unsuitable for use in the subject.

51. A method of determining the presence in a subject of an AAV4-specific antibody comprising administering to an antibody-containing sample from the subject an antigenic fragment of the protein of claim 37 and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the presence of an AAV4-specific antibody in the subject.

52. A method of delivering a nucleic acid to a cell comprising administering to the cell an AAV4 particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell.

53. The method of claim 52, wherein the AAV inverted terminal repeats are AAV4 inverted terminal repeats.

54. The method of claim 52, wherein the AAV inverted terminal repeats are AAV2 inverted terminal repeats.

55. A method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

56. A method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV4 particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject.

57. A method of delivering a nucleic acid to a cell in a subject having antibodies to AAV2 comprising administering to the subject an AAV4 particle comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject.

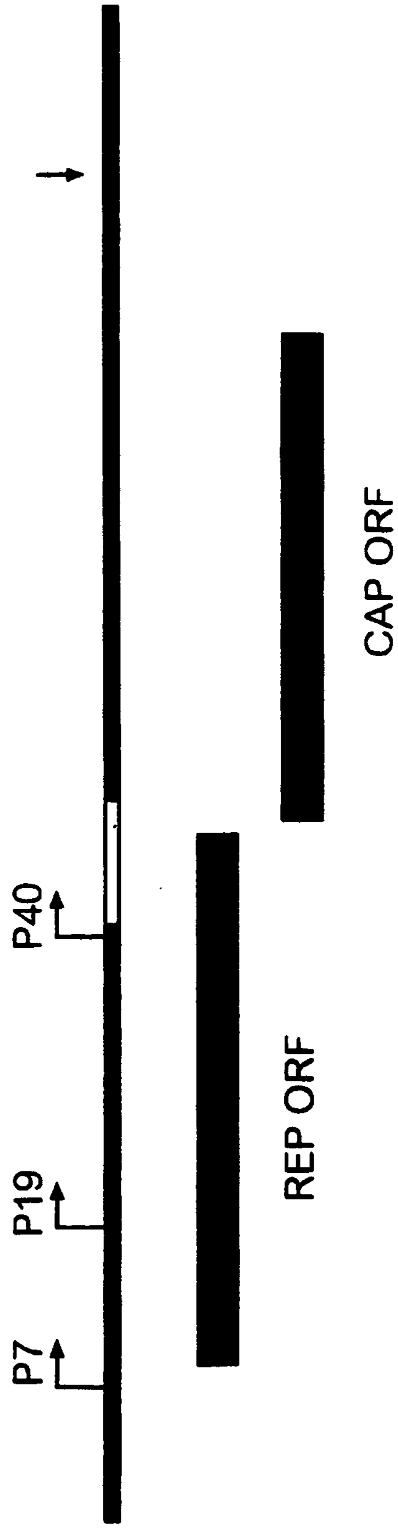


FIG.1

SUBSTITUTE SHEET (RULE 26)

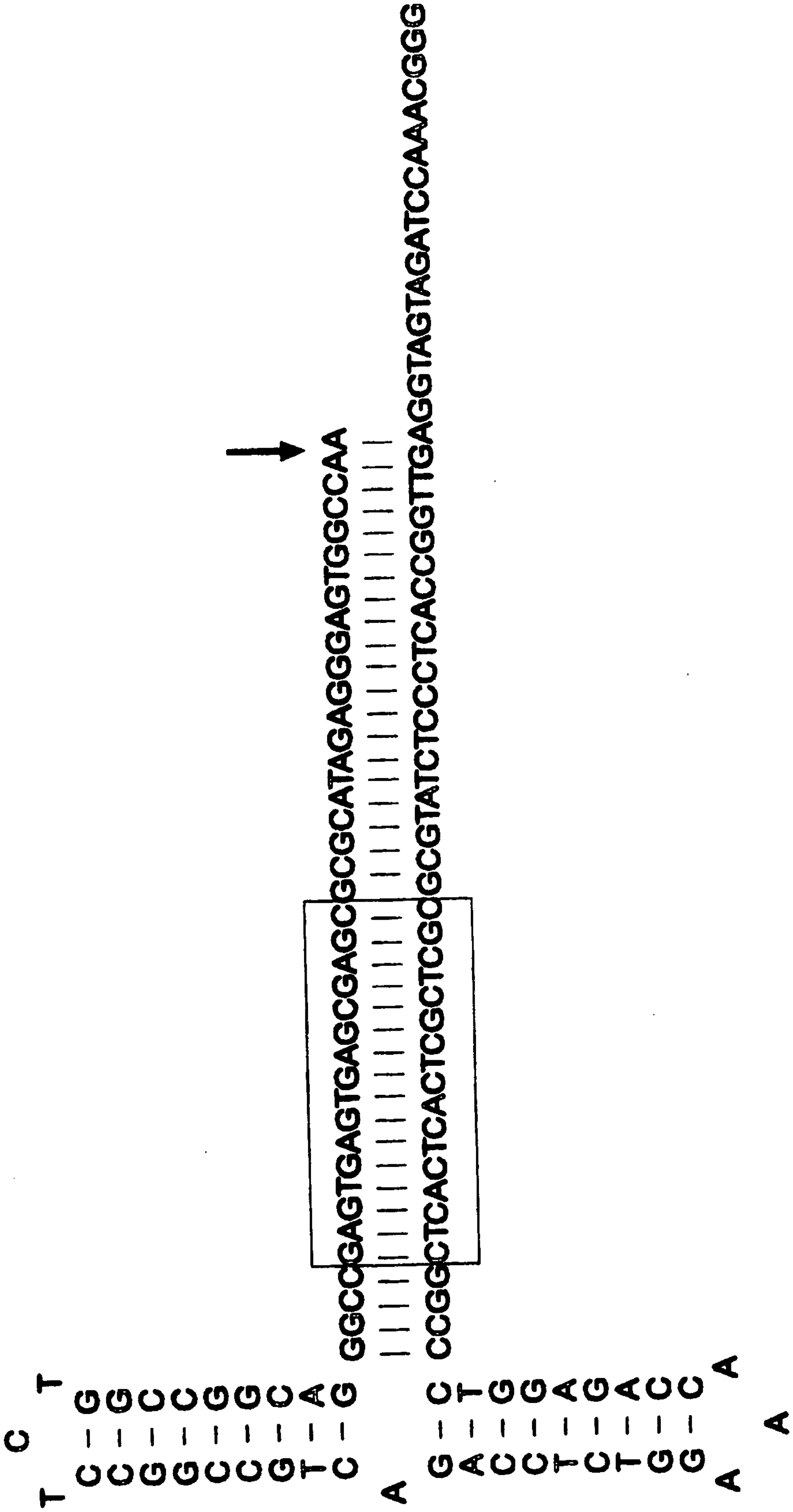

GGCCGAGTGAGTGAGCGAGCGCGCATAGAGGGAGTGGCCAA
CCGGCTCACTCACTCGCTCGCGCGTATCTCCCTCACCGGTTGAGGTAGTAGATCCAAAACGGG

FIG.2

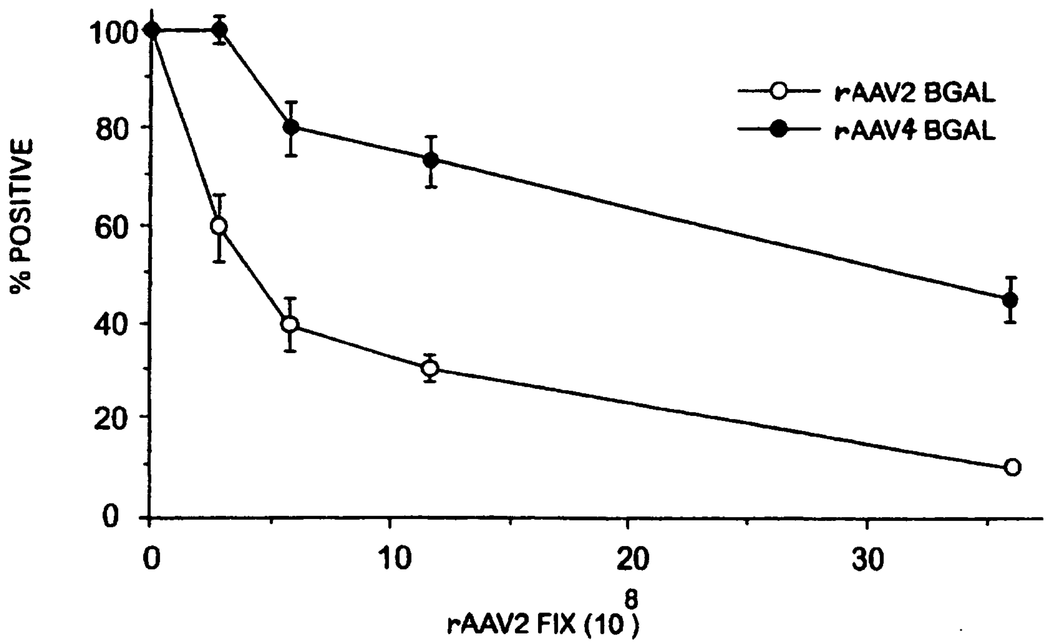

100
80
60
40
20
0
% POSITIVE
0
10
20
30
rAAV2 FIX (10^8)
rAAV2 BGAL
rAAV4 BGAL

FIG.3

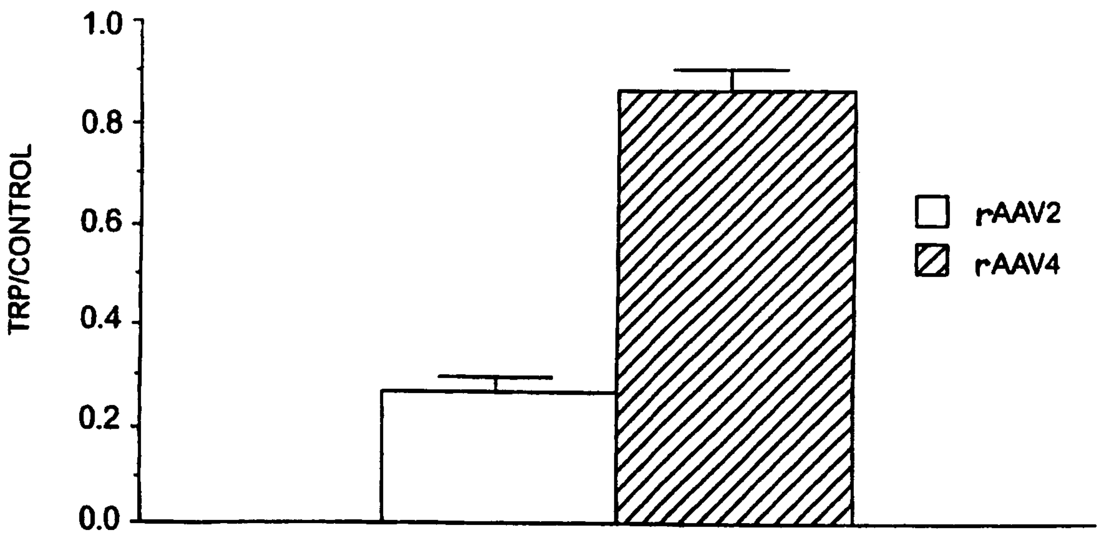

1.0
0.8
0.6
0.4
0.2
0.0
TRP/CONTROL
rAAV2
rAAV4

FIG.4

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/86, C07K 14/015, G01N 33/53, A61K 48/00	A3	(11) International Publication Number: WO 98/11244 (43) International Publication Date: 19 March 1998 (19.03.98)

(21) International Application Number: PCT/US97/16266

(22) International Filing Date: 11 September 1997 (11.09.97)

(30) Priority Data:
60/025,934 11 September 1996 (11.09.96) US

(71) Applicant *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; National Institutes of Health, Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CHIORINI, John, A. [US/US]; 2604 Loma Street, Silver Spring, MD 20902 (US). KOTIN, Robert, M. [US/US]; 707 Gormley, Rockville, MD 20850 (US). SAFER, Brian [US/US]; 1610 Tilton Drive, Silver Springs, MD 20902 (US).

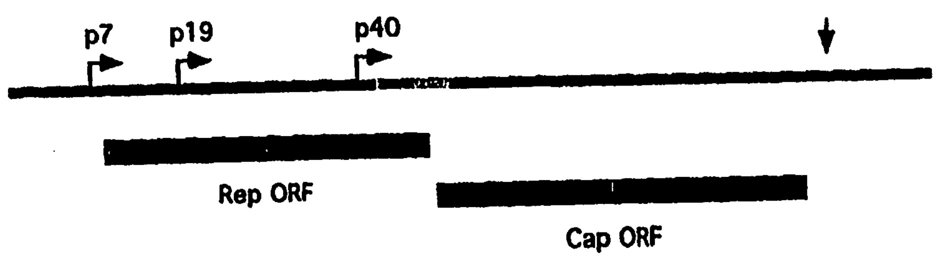
(74) Agents: SELBY, Elizabeth et al.; Needle & Rosenberg, Suite 1200, 127 Peachtree Street, N.E., Atlanta, GA 30303 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 22 May 1998 (22.05.98)

(54) Title: AAV4 VECTOR AND USES THEREOF



(57) Abstract

The present invention provides an adeno-associated virus 4 (AAV4) virus and vectors and particles derived therefrom. In addition, the present invention provides methods of delivering a nucleic acid to a cell using the AAV4 vectors and particles.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

Internatl Application No
PCT/US 97/16266

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/86 C07K14/015 G01N33/53 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 96 18727 A (AVIGEN INC) 20 June 1996 see page 9, paragraph 2 - paragraph 4 see page 10, paragraph 2 - paragraph 3 see page 11, line 13 - line 27 see page 32, line 5 - line 26; figure 7B; examples 1,2 ---	1,10,15, 33,47, 48,52-54
Y	SHIN-ICHI MURAMATSU ET AL: "NUCLEOTIDE SEQUENCING AND GENERATION OF AN INFECTIOUS CLONE OF ADENO-ASSOCIATED VIRUS 3" VIROLOGY, vol. 221, no. 1, 1 July 1996, pages 208-217, XP000608965 see the whole document --- -/--	1,10,15, 33,40, 47,48, 52-54

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

3

Date of the actual completion of the international search	Date of mailing of the international search report
12 March 1998	27.03.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gurdjian, D

INTERNATIONAL SEARCH REPORT

Internati Application No
PCT/US 97/16266

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MUSTER CJ ET AL: "Physical mapping of adeno-associated virus serotype 4 DNA." J VIROL, SEP 1980, 35 (3) P653-61, UNITED STATES, XP002058632 see the whole document	1,10,15, 33,40, 47,48, 52-54
Y	WO 96 00587 A (UNIV PITTSBURGH) 11 January 1996 see page 9, paragraph 2	40
A	SRIVASTAVA A ET AL: "Nucleotide sequence and organization of the adeno-associated virus 2 genome." J VIROL, FEB 1983, 45 (2) P555-64, UNITED STATES, XP002058633 see the whole document	1,10,15, 33,47, 48,52-54
A	SALO R.J. ET AL: "Structural polypeptides of parvoviruses" VIROLOGY, 1977, 78/1 (340-345), USA, XP002058634 see the whole document	1,10,15, 33,47, 52-54
P,X	CHIORINI JA ET AL: "Cloning of adeno-associated virus type 4 (AAV4) and generation of recombinant AAV4 particles." J VIROL, SEP 1997, 71 (9) P6823-33, UNITED STATES, XP002058635 see the whole document	1-46

3

INTERNATIONAL SEARCH REPORT

Inte. .ional application No.
PCT/US 97/16266

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Remark : Although claims 55-57 and 52-54, as far as they concern an in vivo method , are directed to a method of treatment of the human/animal body , the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internati Application No
PCT/US 97/16266

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9618727 A	20-06-96	EP 0793713 A	10-09-97
WO 9600587 A	11-01-96	AU 2913895 A	25-01-96
		CA 2193802 A	11-01-96
		EP 0766569 A	09-04-97